

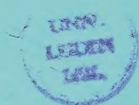
Aus der Psychiatrischen Klinik mit Poliklinik
der Universität Erlangen-Nürnberg
Direktor: Prof. Dr. E. Lungershausen

Prüfung des Syndrom-Kurztests auf
seine Retestreliaibilität an 120
durchschnittlich intelligenten
Patienten

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

vorgelegt von
Gunther Salomo
aus Nürnberg

1983



Aus der Psychiatrischen Klinik mit Poliklinik
der Universität Erlangen-Nürnberg
Direktor: Prof. Dr. E. Lungershausen

Prüfung des Syndrom-Kurztests auf
seine Retestreliabilität an 120
durchschnittlich intelligenten
Patienten

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

vorgelegt von
Gunther Salomo
aus Nürnberg

Gedruckt mit Erlaubnis der Medizinischen Fakultät
der Universität Erlangen-Nürnberg

Dekan: Prof. Dr. F. Wolf
Referent: Prof. Dr. W. Kinzel
Korreferent: Prof. Dr. E. Lungershausen

Tag der mündlichen Prüfung: 31.10.1983

Inhaltsverzeichnis

- I Einleitung
 - 1. Vorbemerkung
 - 2. Fragestellung

- II Allgemeine Theorie
 - 1. Geschichte der Psychopathometrie
 - 2. Theorie des Durchgangs-Syndroms
 - 3. Psychopathometrische Tests
 - 4. Meßtheoretische Probleme
 - 5. Verfahren zur Messung der Funktionspsychose

- III Vorstellung der in den folgenden Untersuchungen
verwendeten Verfahren
 - 1. Der MWT-B
 - 2. Beschreibung des SKT
 - 2.1. Allgemeine Bemerkungen
 - 2.2. Die Einzelaufgaben
 - 3. Güterwerte des SKT

- IV Testtheorie
 - 1. Definition von "Test"
 - 2. Hauptgütekriterien eines Tests
 - 3. Die Axiome der klassischen Testtheorie
 - 4. Fehlerquellen einer Messung

- 5. Berechnung der Reliabilität
 - 5.1. Die Retestmethode
 - 5.2. Die Paralleltestmethode
 - 5.3. Die Methode der Testhalbierung
 - 5.4. Reliabilitätsschätzungen mit Hilfe des "Koeffizienten Alpha"

- V Testdurchführung und Befunde
 - 1. Durchführung der Untersuchung
 - 2. Beschreibung der Untersuchungsstichprobe

- VI Ergebnisse
 - 1. Reliabilitätsberechnungen
 - 2. Korrelationen
 - 3. Einflüsse durch Alter und Intelligenz
 - 4. Positionseffekte

- VII Diskussion

- VIII Zusammenfassung

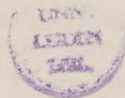
- IX Literaturverzeichnis

I Einleitung

1. Vorbemerkung

Der Terminus "Psychopathometrie" wurde von H.H. WIECK und K.H. STÄCKER im Jahre 1964 eingeführt (50). Die exakte Begriffsbestimmung und wissenschaftstheoretische Begründung erfolgte 10 Jahre später durch W. KINZEL (23). "Psychopathometrie" meint demzufolge "die objektiv verbindliche Messung der Störungsgrade psychopathologischer Erscheinungen" (S.10).

Die erste Vorform eines psychopathometrischen Tests war der 1959 und 1969 von F. BÖCKER vorgestellte Syndrom-Test zur Erfassung von Bewußtseinstörungen und Durchgangs-Syndromen (2;3). Auf diesen aufbauend, legte H. ERZIGKEIT im Jahre 1977 den Syndrom-Kurztest (SKT) vor, der im Rahmen körperlich begründbarer Psychosen Aufmerksamkeits- und Gedächtnisstörungen zu beurteilen gestattet (10). Der SKT hat sich in der Zwischenzeit als ein in Klinik und Praxis einfach zu handhabendes und robustes Verfahren zur Messung der Schweregrade von Durchgangs-Syndromen bewährt (26). Zusammen mit der Funktionspsychose-Skala B, die stärkere Störungen besser erfaßt, sind heute seine Hauptanwendungsgebiete Verlaufsbeobachtungen von Durchgangs-Syndromen sowie die klinische Prüfung von Psychopharmaka (33).



2. Fragestellung

Um durch Lerneffekte hervorgerufene fehlerhafte Ergebnisse bei Wiederholungstestungen, wie sie im Rahmen von Verlaufskontrollen möglich sind, auszuschließen, wurden zusätzlich zu der bereits vorhandenen und normierten Testform A des SKT (13) noch die Testformen B bis E erstellt. Als Voraussetzung gilt, daß alle fünf Testformen äquivalent sind. Nur so können sie ihre Aufgabe erfüllen. Diese besteht darin, bei veränderter Merkmalsausprägung (d.h. bei geringerem bzw. stärkerem Schweregrad der Funktionspsychose, vgl. zweiter Abschnitt des nächsten Kapitels) den jeweils augenblicklichen Störungsgrad möglichst genau zu ermitteln.

Auf die Reliabilität des SKT, d.h. auf den Grad seiner Meßgenauigkeit (zu näheren Erläuterungen des Begriffs vgl. Kapitel IV), liegen bis jetzt nur indirekte Hinweise aus Validitätsuntersuchungen vor (13). Damit ergibt sich die Forderung, eine Reliabilitätsuntersuchung an Hand eines größeren Probandenkollektivs vorzunehmen. Die Reliabilitäten der Testform A sollen dabei über die Berechnung der Retestreliabilität und des Koeffizienten Alpha (nach L.J. CRONBACH) abgeschätzt werden (vgl. dazu Kapitel IV).

Als Nebenaufgaben werden weiterhin der Einfluß von Intelligenz und Alter der Patienten auf die Ergebnisse der fünf Testformen zu untersuchen sein. Ebenso soll geprüft werden, ob die Reihenfolge der Darbietung der einzelnen Testformen bei der Abnahme der Verfahren Auswirkungen auf die Ergebnisse hat.

II Allgemeine Theorie

1. Geschichte der Psychopathometrie

Die ersten Versuche der Erstellung von Verfahren zur Erfassung psychopathologischer Größen (und z.T. auch deren Ausprägung) wurden bereits gegen Ende des vorigen Jahrhunderts unternommen. So stellte R. SOMMER (vgl.1) die Forderung nach vom jeweiligen Untersucher oder Beobachter unabhängigen (also objektiven) Methoden zur Diagnosestellung auf. Phänomene und Störungen im Bereich der geistigen Leistungsfähigkeit sollten somit in eine Norm gebracht und damit vergleichbar werden. Auch E. KRAEPELIN (vgl.1) beklagte das Fehlen objektiver Methoden zur Beurteilung von Störungen im psychischen Bereich. Einen ersten Schritt aus dieser Lage stellte das 1885 von C. RIEGER entwickelte Verfahren zur Prüfung des Schwachsinnnes dar (vgl.48). Auch Namen wie G. ROSSOLIMO, F. GALTON und E. RODENWALDT sind auf das Engste mit jenen Meßansätzen um die Jahrhundertwende verknüpft (vgl.1). Wurde die neue sog. Kurvenpsychiatrie, die von P. NÄCKE (vgl.50) als rationelle und intensive Anwendung der Statistik auf Grund alter und neuer Untersuchungsmethoden verstanden wurde, auch von vielen Psychiatern abgelehnt, so gelang es doch im Laufe der Zeit, Eindrucks- und Kurvenpsychiatrie sinnvoll aufeinander zu beziehen. Gingen die Impulse anfangs von den Psychiatern aus (1), so verlagerte sich die Testentwicklung nach Einführung der apparativen Diagnostik, z.B. des EEG in der Psychiatrie, später in die Hände der Psychologen.

2. Theorie des Durchgangs-Syndroms

Im Jahre 1956 wurde von H.H. WIECK anlässlich der 72. Wanderversammlung südwestdeutscher Neurologen und Psychiater in Baden-Baden erstmals das Durchgangs-Syndrom beschrieben und seine Stellung innerhalb der reversiblen körperlich begründbaren Psychosen angegeben (42;48). Die Durchgangs-Syndrome umfassen die leichteren Störungsgrade der später (1967) von H.H. WIECK so genannten "Funktionspsychose" (45), die das reversible körperlich begründbare Gesamtsyndrom von seinen leichtesten intraindividuellen Normabweichungen bis zur Bewußtlosigkeit umgreift.

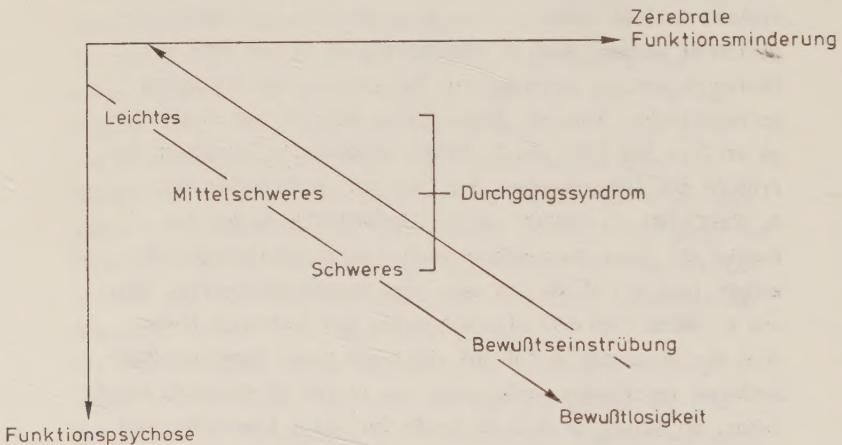


Abb. II.1. Schweregrade der Funktionspsychose

Die psychischen Defizite einer Funktionspsychose sind kausal begründet durch primäre oder sekundäre Hirnbeeinträchtigungen, wie sie beispielsweise Hirnkontusionen, Barbituratver-

giftungen, aber auch die funikuläre Spinalerkrankung und die eklampthische Urämie darstellen können. Die Funktionspsychose ist also nicht als nosologische Diagnose zu verstehen, sondern als Bezeichnung für ein unspezifisches Krankheitsbild, dem verschiedene Hirnerkrankungen zugrunde liegen können (15). "Funktion" soll andeuten, daß dieses Krankheitsbild keine bleibende (irreversible) Veränderung zentraler psychischer Strukturen bezeichnet, sondern eine grundsätzlich rückbildungsfähige (reversible) Störung peripherer psychischer Funktionen. Zum Problem der Trennung von reversiblen und irreversiblen Syndrom bei körperlich begründbaren Psychosen vgl. W. KINZEL (20;21;22;24).

Die sog. symptomatischen Psychosen (heute: körperlich begründbare Psychosen) wurden in den Jahren 1910 und 1917 von K. BONHOEFFER (4;5) als unspezifisch beschrieben, das heißt: Verschiedene (Gehirn-) Krankheiten können zu den gleichen psychopathologischen Bildern führen. Ebenso vermögen alle möglichen psychopathologischen Prägnanztypen symptomatischer Psychosen bei nur einer Grundkrankheit aufzutreten.

H.H. WIECK hat dieses Prinzip für das reversible Syndrom um ein weiteres, das der quantitativen Kausalität, erweitert. Es besagt, daß die Schwere der zerebralen Funktionsstörung das Ausmaß der psychopathologischen Beeinträchtigung bestimmt. Gelingt es also, typische (Leistungs-) Defizite einer Funktionspsychose, zum Beispiel Aufmerksamkeits- und Gedächtnisstörungen, quantifizierend zu erfassen, so sind damit auch Aussagen über den Schweregrad der zugrunde liegenden Hirnbeeinträchtigung möglich (15;48).

Nach H.H. WIECK (43;44;49) unterliegt die Funktionspsychose dem "Prinzip der homogenen Syndrondynamik", demzufolge alle Einzelfunktionen in jeweils gleicher Ausprägungsstärke

betroffen sein sollen. Es zeigt sich jedoch, daß die Objektivierung dieser Hypothese methodisch noch nicht befriedigend gelungen ist (32).

Grundlage der Meßbarkeit von Schweregraden der Funktionspsychose ist der Umstand, daß die "Achsensymptomatik" des reversiblen Syndroms die "Minusleistungen" sind (die sog. produktiven Symptome = die "Plussymptome" gelten als "fakultative Ausgestaltungen"). Auf diese Weise ergeben sich die bereits weiter vorne (vgl. Abb. II.1) dargestellten Einzelstadien der Funktionspsychose von den leichtesten Störungsgraden des Durchgangs-Syndroms bis zur Bewußtlosigkeit (vgl. auch 20 u. 26).

3. Psychopathometrische Tests

Psychometrische Tests beschäftigen sich mit der Messung von psychischen Funktionen und Merkmalen im normalpsychologischen Bereich, während psychopathometrische Verfahren die Ausprägungsgrade der psychopathologischen Störungen objektiv zu erfassen suchen (23;48). Psychometrie setzt also den "Normalzustand" voraus, damit im wesentlichen Merkmale von relativer zeitlicher Konstanz. Psychopathometrie bei Funktionspsychosen hingegen geht von einer Leistungseinschränkung aus, also von Merkmalen, die sich pro Zeiteinheit in ihrer quantitativen Ausprägung rasch ändern können. Man spricht deshalb auch von der "Dynamik" einer Funktionspsychose (20).

Zu den Aufgaben psychopathometrischer Forschungen gehört es, Aussagen über Art und Grad der jeweiligen Störung zu machen sowie im Rahmen von Verlaufskontrollen die Verlaufsrichtung des Prozesses ("Besserung" oder "Verschlechterung") zu bestimmen. Vorab bei therapeutischen Maßnahmen ist es wertvoll, Erfolge nicht nur durch die "Erfahrung" des Therapeuten, sondern auch durch Methoden, die eine objektive Überprüfbarkeit gestatten, zu ermitteln.

Gegenüber anderen sog. klassischen Verfahren wie beispielsweise dem Hamburg-WECHSLER-Intelligenztest für Erwachsene (HAWIE) oder dem RORSCHACH-Test unterscheiden sich psychopathometrische Verfahren durch ihre spezifische Angemessenheit an den jeweiligen psychopathologischen Untersuchungsgegenstand. So werden psychopathometrische Tests i.a. kurz sein, um eine geringere Belastung der Patienten zu gewährleisten (51).

Fragebogenverfahren wie Selbst- oder Fremdbeurteilungsskalen sind in der Durchführung sehr ökonomisch, bedürfen einer

meist nur kurzen Einarbeitungszeit und zeigen sich den meisten Störeinflüssen gegenüber wenig anfällig (51). Die Anwendung von Selbstbeurteilungsverfahren ist nur dann sinnvoll, wenn man eine mindestens durchschnittliche Intelligenz ($HAWIE-IQ \geq 91$) voraussetzen kann (28). Bei Fremdbeurteilungsverfahren schätzen Arzt oder medizinisches Personal den jeweiligen Zustand des Patienten ein. Diese Methoden sind aber auf Grund ihrer Konstruktion nur selten befriedigend valide und reliabel (12). Der Nachteil von Fragebogenverfahren im psychopathologischen Bereich liegt in erster Linie darin begründet, daß schon von mittelschweren Störungsgraden an die Ergebnisse nicht mehr verwertbar sind (51).

Hier zeigen sich Leistungsverfahren als ergiebiger. Sie erfassen schwere Störungsgrade auch im Falle zylothymen Depressionen beispielsweise wesentlich besser als Selbstbeurteilungsbögen (9;51).

Selbstverständlich ist -zumindest zunächst- auch im Bereich psychopathometrischer Messungen eine standardisierte Test-erhebung, die den Ablauf der Testung genau festlegt, notwendige Bedingung für valide und reliable Testbefunde.

4. Meßtheoretische Probleme

Da es sich bei der Funktionspsychose um Veränderungen im seelisch-geistigen Innenbereich handelt, das Störungsausmaß aber nur von außen bestimmt werden kann, findet der Störungsgrad nicht unbeeinflusst seinen Niederschlag im Testergebnis. Es müssen also Moderatorvariable, Störvariable und Artefakte beachtet werden (48).

Als Moderatorvariablen bezeichnet man jene Größen, die das gemessene Merkmal in seiner Ausprägung verändern können. Als wichtigste sind im Zusammenhang mit der Messung des Durchgangs-Syndroms die Variablen "Alter" und "Intelligenz" zu nennen (51). Sowohl bei Leistungstests als auch in großem Umfange bei Fragebogenverfahren ist die Intelligenz beinahe prinzipiell als Moderatorvariable zu werten. Um den Einfluß dieser Größe so gering wie möglich zu halten, sollten bei guten Tests die Moderatorvariablen mit in die Normierung eingehen. Einflußgrößen, die unkontrolliert das gemessene Merkmal beeinflussen, nennt man Störvariablen. Lassen sich äussere Störgrößen, wie eine unbequeme Sitzposition, schlechte Beleuchtung und Lärm durch Wahl des Testraumes noch beseitigen, so sind bei schwerkranken Patienten Einflüsse durch pflegerische Verrichtungen, Dauerinfusionen, Medikamente oder Bewegungseinschränkung nicht vollständig eliminierbar. Tageszeit, Motivation, emotionale Lage des Patienten (z.B. nach Ankündigung einer Operation oder Verkündung der Diagnose) dürfen auch nicht vernachlässigt werden. Mag eine Ersttestung für den Patienten noch interessant und abwechslungsreich sein, so werden Wiederholungen doch schon häufiger als Belastung empfunden. Dies zumal dann, wenn deren Sinn nicht unmittelbar einzusehen ist (51). Außerdem muß berücksichtigt werden,

daß kranke Menschen oft nur kurze Zeit in der Lage sind, eine Tätigkeit konzentriert durchzuführen. Bei einer Anwendung von längeren Leistungsverfahren würde man dann in erster Linie so etwas wie die "Durchhaltefähigkeit" prüfen (was man ja nicht will, da eine Aussage über diese "Eigenschaft" nicht Ziel der Untersuchung ist). Als weitere Komponenten, die das wahre Testergebnis verfälschend beeinflussen können, sind noch Versuchsleitereinflüsse, allgemeine Beurteilungsfehler wie Halo-Effekt oder ROSENTHAL-Effekt und die Vigilanz des Patienten zu nennen. Bei Erkrankungen mit hoher Merkmalspenetranz, wie bei Funktionspsychosen höherer Störungsgrade, wirken sich die meisten Störgrößen jedoch verständlicherweise geringer aus (48).

Artefakte können entstehen, wenn ein Patient bei einer Wiederholungstestung durch Lerneffekte oder einfaches Wiedererkennen der Aufgaben ein besseres Ergebnis erzielt, obwohl sein Leistungsniveau "in Wahrheit" dasselbe geblieben ist. Durch Darbietung von Parallelförmigkeiten oder durch einen ausreichenden zeitlichen Abstand der Testungen versucht man, diesen Lerneffekt, der mit zunehmenden Störungsgrad naturgemäß immer weniger in Erscheinung tritt, so gering wie möglich zu halten. Allerdings können auch neuropsychologische Beeinträchtigungen zu erheblichen Artefakten führen, ohne daß deren Einfluß bisher völlig ausgeschaltet werden kann.

5. Verfahren zur Messung der Funktionspsychose

Das erste brauchbare Ergebnis in dem Bemühen, die theoretischen Neuerungen durch empirische Methoden zu ergänzen, stellt der von F. BÖCKER (2;3) entwickelte und nach ihm benannte Syndromtest dar. Es war der erste Vorläufer eines psychopathometrischen Leistungsverfahren mit dem Ziel der Schweregradmessung einer (reversiblen, körperlich begründbaren) psychischen Leistungsstörung. Das Verfahren selbst bestand aus dreizehn immer in der gleichen Reihenfolge dargebotenen Aufgaben. Die Lösungszeit wurde mit der Stoppuhr gemessen und die jeweilige Leistungsminderung mit 0 bis 3 Minuspunkten pro Einzelaufgabe bewertet. Bei schwersten Leistungsdefiziten ergaben sich weitere (negative) Zusatzpunkte, je nachdem, ob der Kranke einfache (wenn auch sinnlose) Sätze oder Worte noch zu artikulieren vermochte oder nicht.

Eine Weiterentwicklung dieses ersten Verfahrens zur Schweregradbestimmung des reversiblen Syndroms stellt der Syndrom-Kurztest (SKT) nach H. ERZIGKEIT dar (11;13). Dieses Meßinstrument berücksichtigt bereits weitestgehend jene Erkenntnisse, die im letzten Abschnitt beschrieben worden sind. Vom mittelschweren Durchgangs-Syndrom an bis zu den noch erheblicheren Schweregraden hin hat sich die Orientierungsskala nach K.F. DRUSCHKY bewährt. Sie enthält Aufgaben zur örtlichen, zeitlichen, personellen und situativen Orientierung stationärer Patienten (7). Die Orientierungsskala wurde zur Funktionspsychose-Skala A weiterentwickelt. Die Skala besteht aus 27 Aufgaben, die in einer ansteigenden Schwierigkeitsreihe geordnet sind. Können fünf oder sieben Aufgaben hintereinander nicht gelöst werden, so wird der

Test abgebrochen (Testökonomie!). Eine nochmalige Erweiterung stellt die Funktionspsychose-Skala B dar, die eine zusätzliche Unterscheidung der Schweregrade der neurologischen Komasympptomatik erlaubt (34).

Komatöse Zustände hingegen lassen sich am besten mit der Munich-Coma-Scale nach R. BRINKMANN, D.v. CRAMON und H. SCHULZ (6) differenzieren. Diese Methode setzt sich aus zwei Skalen zusammen, von denen die eine zur Abstufung der Stimulierbarkeit, die andere zur Unterscheidung der Reaktivität des Patienten dient.

vornehmlich leichtes und mittelschweres Durchgangs-Syndrom	Syndromtest nach F. BÖCKER SKT nach H. ERZIGKEIT
vornehmlich mittel- schweres und schweres Durchgangs-Syndrom	Orientierungsskala nach K.F. DRUSCHKY Funktionspsychose-Skala B
komatöse Zustandsbilder	MCS nach R. BRINKMANN, D.v. CRAMON u. H. SCHULZ

Abb. II.2. Verfahren zur Messung der Funktionspsychose

III Vorstellung der in den folgenden

Untersuchungen verwendeten Verfahren

1. Der MWT - B

Wie schon erwähnt, stellt die Intelligenz bei Leistungsverfahren eine prinzipielle Moderatorvariable dar. Dieser Sachverhalt läßt sich besonders eindrucksvoll am Beispiel des Durchgangs-Syndroms demonstrieren (51). Während der leichten Stadien dieser Phase des reversiblen Syndroms erzielt der Intelligentere in einem Leistungstest immer noch ein besseres Ergebnis als ein weniger begabter Mensch ohne Durchgangs-Syndrom. Würde man, um die Intelligenz zu bestimmen, einen der üblichen Intelligenztests anwenden, so wäre das Ergebnis dieser Messung vom Durchgangs-Syndrom beeinflußt und damit nur bedingt interpretierbar. Man könnte, nachdem die Zusammenhänge untersucht sind (40), bei Kenntnis des Schweregrades des reversiblen Syndroms und des HAWIE-Intelligenzquotienten auf den prämorbid Intelligenzquotienten "zurückschliessen" oder durch andere Informationen, etwa durch Berücksichtigung des Berufs und der Schulbildung, zu einer Entscheidung gelangen (40). Diese Verfahrensweisen sind aber sehr umständlich und ungenau. Man hat deshalb einen Test konstruiert, der in kurzer Zeit, auch bei Patienten mit Störungsgraden, die das mittelschwere Durchgangs-Syndrom einbeziehen, eine genügend genaue Schätzung des prämorbid allgemeinen Intelligenzniveaus ermöglicht. Dieses Verfahren stellt der MWT-B (27), eine Weiterentwicklung des von S. LEHRL et al. 1971 eingeführten Mehrfachwahl-

-Wortschatz-Tests (30), dar. Von allen Varianten dieses Instrumentariums ist der MWT-B das am aufwendigsten geeichte Verfahren (31;38). Standardnormen wurden an einer repräsentativen Stichprobe deutschsprachiger Erwachsener ($n = 1952$) ermittelt (29). Der Test kann überall dort durchgeführt werden, wo ein aufwendiges Intelligenztestverfahren nicht zuzumuten ist. Der MWT-B hat also die Eigenschaft, gegenüber psychischen Störeinflüssen -wie auch immer bedingt- relativ unanfällig zu sein. Für den hier vorgesehenen Anwendungsbereich mißt er hinreichend genau. Genügt eine grobe Schätzung, so vermag er auch auf der "Normalebene" gute Dienste zu leisten. Beim MWT-B gilt es, aus einer Reihe von fünf Worten, von denen vier wissenschaftlich und umgangssprachlich nicht existieren, das einzig richtige, d.h. im Duden vorfindbare Wort, herauszufinden. Dargeboten werden 37 Wortreihen ansteigenden Schwierigkeitsgrades, wobei 21 - 30 richtige Antworten der "durchschnittlichen Intelligenz" im Sinne des HAWIE (nach D. WECHSLER) entsprechen (vgl. Abbildung III.1 und III.2 auf den nächsten Seiten).

Die Zufallswahrscheinlichkeit eines "Treffers" ist bei fünf Möglichkeiten mit 20% sehr hoch. Die Befunde zur Reliabilität und Validität zeigen dennoch recht gute Ergebnisse (38).

Anweisung:

Sie sehen hier mehrere Reihen mit Wörtern. In jeder Reihe steht höchstens ein Wort, das Ihnen vielleicht bekannt ist. Wenn Sie es gefunden haben, streichen Sie es bitte durch.

1. Nale – Sahe – Nase – Nesa – Sehna
2. Funktion – Kunition – Finzahn – Tuntion – Tunkion
3. Struk – Streik – Sturk – Strek – Kreik
4. Kulisse – Kulerane – Kulisse – Klubhile – Kubstana
5. Kenekel – Gesock – Kelume – Gelenk – Gellerg
6. siziol – satzahl – sozihl – sziam – sozial
7. Sympasie – Symmoletrie – Symmantrie – Symphonile – Symplanie
8. Umna – Pamme – Nello – Ampe – Amme
9. Kusse – Surke – Krustelle – Kruste – Struka
10. Kirse – Sike – Krise – Krospe – Serie
11. Tinxur – Kikukur – Fraktan – Tinkur – Rimsuhr
12. Urfusion – Fudision – Infusion – Syntusion – Nuridion
13. Feudasmus – Fondersmus – Föderalismus – Födismus – Föderasmus
14. Reodor – Radium – Terion – Damin – Oraklum
15. kentern – knerte – kanzen – kretern – kretern

16. Kantate – Rakante – Kenture – Kruthe – Kallara
17. schallern – waschieren – wakieren – schackieren – kaschieren
18. Tuhl – Lar – Lest – Dall – Ud
19. Dissonanz – Diskrisanz – Disfranz – Diholanz – Siolenz
20. Ferindo – Infirno – Orfina – Franelto – Infindio
21. Riklase – Kilster – Rilker – Kistler – Linkure
22. kuhnesisch – kulinarisch – kumensisch – kulissarisch – kamastrisch
23. Roato – Torso – Soro – Torgos – Tosor
24. Kiebar – Beikel – Kelbel – Reiker – Blekerl
25. Raik – Korre – Ruckee – Recke – Ulte
26. Lamone – Talane – Matrone – Tarone – Malorite
27. Turna – Unet – Maut – Taum – Muta
28. Sorekin – Sarowin – Rosakin – Narosin – Kerosin
29. beralen – gerälet – anälteren – unlären – verblären
30. Kapaun – Paukan – Naupack – Aupeck – Ankepan
31. Schaber – Bassiker – Kassiber – Sassiker – Askber
32. Pucker – Keuper – Eucker – Reuspeck – Urkane
33. Sprine – Saprin – Parsin – Purin – Asprint
34. Kulon – Solgun – Koskan – Soran – Klonus
35. Adept – Padet – Edapt – Epalt – Tapod
36. Girdelat – Tingerat – indigenat – Nitgesaar – Ringelaar
37. Berkizie – Brekie – Brakize – Brikazie – Bakria

Abb. III.1

Die 37 dargebotenen Wortreihen des MWT-B

MWT-B: Normen von 1952 Erwachsenen der BRD (repräsentative Zufallsstichprobe)

Rohwert	kumulierter Prozentrang	Standardwert	IQ
0	0.8	74	61
1	0.9	76	64
2	0.9	77	65
3	1.1	77	66
4	1.5	78	67
5	3.4	80	70
6	4.5	82	73
7	4.9	83	74
8	5.3	84	76
9	6.1	84	76
10	6.8	85	77
11	7.7	85	78
12	8.5	86	79
13	9.7	87	80
14	10.9	87	81
15	13.1	88	82
16	15.0	89	83
17	17.2	90	85
18	19.3	91	86
19	22.1	92	88
20	24.7	93	89
21	27.6	94	91
22	30.8	95	92
23	34.2	95	93
24	37.4	96	94
25	41.0	97	95
26	46.1	98	97
27	51.9	100	100
28	58.2	101	101
29	66.1	103	104
30	76.2	105	107
31	85.1	108	112
32	92.2	112	118
33	96.5	116	124
34	98.8	120	130
35	99.7	124	136
36	99.9	129	143
37	100.0	130	145

Abb. III.2

Umrechnungstabelle der Rohwerte des MWT-B

2. Beschreibung des Syndrom-Kurztests

2.1. Allgemeine Bemerkungen

Der SKT als Leistungsverfahren mißt ein breites Spektrum psychischer Funktionen, in erster Linie jedoch Aufmerksamkeits und Gedächtnisleistungen. Da bei Leistungstests Alter und Intelligenz eine gewichtige Rolle spielen, wurden zur Interpretation der Testergebnisse Normen erstellt, bei denen Lebensalter und prämorbid Intelligenzniveau gesondert berücksichtigt sind. Der Geltungsbereich des Tests erstreckt sich von 17- bis "über 64-jährige". Bei der Konzeption der Testaufgaben achtete man darauf, daß die Bearbeitung auch von motorisch weniger geschickten oder gar leicht behinderten liegenden Patienten noch ohne Leistungseinbußen möglich ist (11;12).

Der SKT ist in fünf Parallelformen erhältlich. Jede Form setzt sich aus neun Subtests zusammen. Die maximale Bearbeitungszeit beträgt pro Subtest eine Minute. Die gesamte Einzeltestung nimmt damit in der Regel nicht mehr als 12 bis 13 Minuten in Anspruch. Bei Subtests, die möglichst schnell bearbeitet werden sollen, wird der Proband ausdrücklich darauf hingewiesen. Die Zeiten werden mit einer Stoppuhr gemessen und sofort in einen Ergebnisbogen (siehe Abb. III.3) eingetragen. Ebenso wird mit den anderen meßbaren Ergebnissen verfahren.

Form A

SYNDROM-KURZTEST

Zur Erfassung von Aufmerksamkeits- und Gedächtnisstörungen



Name: _____ Geburtsdatum: _____ Datum: _____
 Beruf: _____ Alter: _____ Uhrzeit: _____
 Diagnose: _____ IQ-Gruppe: _____
 Bemerkungen: _____
 Versuchsleiter: _____

SUBTESTS (Höchstzeit 60 Sekunden)		Rohpunkte	Wert- punkte *	Konfabu- lationen											
I	Gegenstände benennen	☐ Sekunden													
II	Gegenstände unmittelbar reproduzieren														
	<table border="1"> <tr> <td>Rube</td> <td>Fahne</td> <td>Ofen</td> <td>Kuh</td> </tr> <tr> <td>Schiff</td> <td>Bett</td> <td>Lok</td> <td>Kreuz</td> </tr> <tr> <td>Kanne</td> <td>Spritze</td> <td>Engel</td> <td>Schaf</td> </tr> </table> ☐ Fehlende	Rube	Fahne		Ofen	Kuh	Schiff	Bett	Lok	Kreuz	Kanne	Spritze	Engel	Schaf	
Rube	Fahne	Ofen	Kuh												
Schiff	Bett	Lok	Kreuz												
Kanne	Spritze	Engel	Schaf												
Gegenstände bitte nochmals 5 Sekunden zeigen															
III	Zahlen lesen	☐ Sekunden													
IV	Zahlen ordnen	☐ Sekunden													
V	Zahlen zurücklegen	☐ Sekunden													
VI	Symbole zählen * ★ ☐ * 44	☐ Sekunden													
VII	Interferenz ABAB Richtige Folge: BABBABAA BBABABAB BBABABAA ABABABAB	☐ Sekunden													
VIII	Gegenstände mittelbar reproduzieren														
	<table border="1"> <tr> <td>Rube</td> <td>Fahne</td> <td>Ofen</td> <td>Kuh</td> </tr> <tr> <td>Schiff</td> <td>Bett</td> <td>Lok</td> <td>Kreuz</td> </tr> <tr> <td>Kanne</td> <td>Spritze</td> <td>Engel</td> <td>Schaf</td> </tr> </table> ☐ Fehlende	Rube	Fahne	Ofen	Kuh	Schiff	Bett	Lok	Kreuz	Kanne	Spritze	Engel	Schaf		
Rube	Fahne	Ofen	Kuh												
Schiff	Bett	Lok	Kreuz												
Kanne	Spritze	Engel	Schaf												
IX	Gegenstände wiedererkennen														
	<table border="1"> <tr> <td>Rube</td> <td>Fahne</td> <td>Ofen</td> <td>Kuh</td> </tr> <tr> <td>Schiff</td> <td>Bett</td> <td>Lok</td> <td>Kreuz</td> </tr> <tr> <td>Kanne</td> <td>Spritze</td> <td>Engel</td> <td>Schaf</td> </tr> </table> ☐ Fehlende	Rube	Fahne	Ofen	Kuh	Schiff	Bett	Lok	Kreuz	Kanne	Spritze	Engel	Schaf		
Rube	Fahne	Ofen	Kuh												
Schiff	Bett	Lok	Kreuz												
Kanne	Spritze	Engel	Schaf												
Gesamtpunkte															

* siehe Normentabelle im Manual.

Abb. III.3

SKT-Ergebnisbogen, hier Form A

2.2. Die Einzelaufgaben

Subtest I : Gegenstände benennen

Dem Patienten wird eine Tafel mit den Abbildungen von zwölf verschiedenen Gegenständen vorgelegt. Aufgabe ist es nun, die Gegenstände so schnell wie möglich zu benennen und sie sich gleichzeitig einzuprägen, da sie anschließend noch einmal abgefragt werden.



Abb. III.4

Vorlage für Subtests I und II

Subtest II : Gegenstände unmittelbar reproduzieren

Der Patient wird aufgefordert, diejenigen Gegenstände aufzuzählen, an die er sich noch erinnern kann. Dabei spielt die Reihenfolge keine Rolle. Es stehen maximal 60 Sekunden zur Verfügung. "Konfabulationen", d.h. die Nennung von Gegenständen die nicht auf der Tafel enthalten waren, werden vermerkt, gehen aber nicht in die Wertung ein. Anschließend zeigt der Versuchsleiter noch einmal die Tafel für fünf Sekunden mit

dem Hinweis, die Gegenstände sich wiederholt einzuprägen, da gegen Ende der Testung wiederum nach ihnen gefragt werden wird.

Subtest III : Zahlen lesen

Diese und die nächsten beiden Aufgaben werden mit Hilfe einer Magnetplatte, auf der zehn bunte, mit zweistelligen Zahlen versehene Klötzchen haften, durchgeführt. Als erstes sollen diese Zahlen so schnell wie möglich vorgelesen werden.

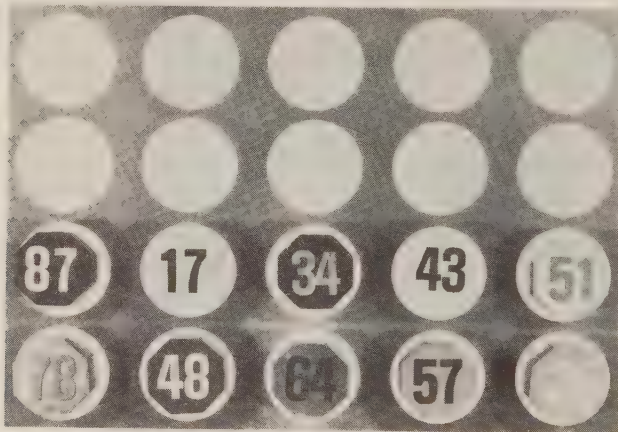


Abb. III.5

Vorlage für Subtests III, IV und V

Subtest IV : Zahlen ordnen

Der Proband hat jetzt die Zahlen wiederum so schnell wie möglich ihrem Zahlenwert nach zu ordnen.

Subtest V : Zahlen zurückstellen

Die Aufgabe besteht hier darin, die Klötzchen so schnell wie

möglich an ihren ursprünglichen Platz zurückzustellen
(Klötzchenziffer zu Unterlageziffer).



Abb. III.6

Vorlage für Subtest V

Subtest VI : Symbole zählen

Auf einer Tafel mit zufällig verteilten Symbolen hat der Patient die Häufigkeit zu bestimmen, mit der eine bestimmte Symbolform ("Kästchen", "Stern" oder "Blume") verzeichnet ist. Nebenbei werden noch Zählfehler vermerkt. Sie bleiben allerdings ohne Konsequenzen auf das Testergebnis (vgl. Abb. III.7 auf der nächsten Seite).

Subtest VII : Interferenzaufgabe

Die auf der Vorlage dargestellten Buchstabenreihen sollen vertauscht gelesen werden, d.h. im vorliegenden Fall: Für A ist ein B und für B ist jeweils ein A laut und so schnell wie möglich zu lesen (vgl. Abb. III.8 auf der nächsten Seite).



Abb. III.7. Vorlage für Subtest VI

ABBABA

ABAABABBAABABABBA
 AABABABBBABAABABA

Abb. III.8. Vorlage für Subtest VII

Subtest VIII : Gegenstände (aus Subtest I) mittelbar
reproduzieren

Der Patient wird aufgefordert, diejenigen Gegenstände aus Subtest I zu nennen, an die er sich jetzt noch erinnern kann. Dazu stehen ihm 60 sec. zur Verfügung, Konfabulationen werden ebenfalls notiert.

Subtest IX : Gegenstände (aus Subtest I) wiedererkennen

Dem Patienten wird eine Tafel mit 48 abgebildeten Gegenständen vorgelegt. Unter ihnen befinden sich auch die 12 Bilder von Subtest I. Diese gilt es jetzt herauszufinden und zu nennen. Die Bearbeitungszeit beträgt wiederum 60 sec. Etwaige Konfabulationen werden wie bei den Aufgaben II und VIII notiert, jedoch nicht bewertet.



Abb. III.9. Vorlage für Subtest IX

Die im Testprotokoll (vgl. Abb. III.3, S.18) in Form von Rohpunkten (RP) eingetragenen Werte werden mit Hilfe von Normtabellen (siehe Tab. V.6a u. b, S.44 u.45) in Wertpunkte (= standardisierte Rohpunkte) umgerechnet. Je nach Abweichung ergeben sich 0 bis 3 Wertpunkte. Dabei sind vier Alters- und drei Intelligenzgruppen zu unterscheiden.

Lebensalter in Jahren			
17 - 44	45 - 54	55 - 64	über 64
IQ - Gruppen			
bis 90	91 - 110	über 110	

Abb. III.10

Einteilung in Alters- und Intelligenzgruppen

Die Summe der Wertpunkte aller neun Untertests erlaubt nun nach folgender Tabelle die Feststellung der Schweregrade der Funktionspsychose.

Gesamtwertpunkte	Schweregrad der Funktionspsychose
0 bis 4	keine Funktionspsychose nachweisbar
5 bis 8	fragliches (oder sehr leichtes) Durchgangs-Syndrom
9 bis 13	leichtes Durchgangs-Syndrom
14 bis 18	mittelschweres Durchgangs-Syndrom
19 bis 23	schweres Durchgangs-Syndrom
24 bis 27	Bewußtseinstörung

Abb. III.11

Tabelle zur Interpretation der Gesamtwertpunkte im SKT

3. Gütwerte des SKT

Normierung:

Die Normierung des SKT erfolgte an 625 ("Normal"-) Personen, die sich auf drei Intelligenz- und vier Altersstufen verteilen.

Validität:

Von der Konzeption des Testes her kann bereits auf eine ausreichende (Konstrukt-) Validität des Verfahrens geschlossen werden. Bei Vergleichen mit anderen Verfahren, die den Anspruch erheben, funktionspsychotische Schweregrade zu messen, ergaben sich folgende Ergebnisse: Die Korrelation verschiedener Leistungsverfahren für Funktionspsychosen zusammen mit dem Gesamtergebnis des Syndrom-Tests beträgt $\rho = 0,71$ (16). Die Korrelation mit der FPS-A, einer Vorgängerin der FPS-B, beträgt $\rho = 0,95$ (36).

Reliabilität:

Reliabilitätsmessungen — sie sind ja Gegenstand dieser Untersuchung — gestalten sich sehr schwierig und wurden bisher nicht durchgeführt. Auf die Höhe der Reliabilität liegen deshalb nur indirekte Hinweise aus Validitätsüberprüfungen vor. Man macht sich dabei die Regel zunutze, daß die Reliabilitätskoeffizienten mindestens so groß sein müssen wie die Wurzel aus den Validitätskoeffizienten (13).

IV Testtheorie

1. Definition von "Test"

Das Wort "Test" kommt aus dem englischen Sprachgebrauch und bedeutet soviel wie Probe. In der Psychologie wird der Testbegriff nach G.A. LIENERT folgendermaßen definiert:

"Ein Test ist ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung." (35,S.7).

2. Hauptgütekriterien eines Tests

Die Hauptgütekriterien eines Tests sind:

Objektivität, Reliabilität und Validität. Daran schließen sich vier Nebenkriterien an. Demzufolge soll ein Test normiert, vergleichbar, ökonomisch und nützlich sein (35). Unter der Objektivität eines Tests versteht man den Grad, in dem die Ergebnisse unabhängig vom Untersucher sind. Die Forderung nach Objektivität bedeutet also, daß verschiedene Untersucher bei demselben Probanden zu gleichen Ergebnissen gelangen. Analog ist eine Aufgabe objektiv, wenn sie von verschiedenen Beurteilern übereinstimmend als richtig bzw. falsch oder kennzeichnend für das Vorhandensein oder Nichtvorhandensein eines Merkmals gewertet wird (35;40).

Unter der Reliabilität eines Tests versteht man den Grad der Genauigkeit, mit dem er ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal mißt, gleichgültig, ob er dieses Merkmal auch zu messen beansprucht. Ein Test wäre also ideal reliabel, wenn die erzielten Ergebnisse den Probanden fehlerfrei beschreiben; eine Aufgabe ist reliabel, wenn bei ihrer Wiederholung nach angemessener Zeit bei unverändertem Persönlichkeitsmerkmal des Probanden die Beantwortung in derselben Weise erfolgt. Der Grad der Reliabilität wird durch einen Koeffizienten ausgedrückt, der die Übereinstimmung zweier unter gleichen Bedingungen gewonnener Meßergebnisse darstellt (35).

Unter der Validität eines Tests versteht man den Grad der Genauigkeit, mit dem der Test dasjenige Persönlichkeitsmerkmal, das er zu messen vorgibt, tatsächlich mißt. Ein Test ist demnach vollkommen valide, wenn seine Ergebnisse einen fehlerfreien Schluß auf das zu erfassende Merkmal einschließlich Ausprägungsgrad zulassen. Eine Aufgabe ist valide, wenn sie von Probanden mit starker Merkmalsausprägung häufiger im Sinne der Erwartung beantwortet wird als von Probanden mit geringer Merkmalsausprägung (35;39).

Die Validität ist definiert als die Korrelation eines Tests mit einem Kriterium (14). Je nach Wahl des Kriteriums kann ein Test über mehrere Validitäten verfügen. Wählt man als Kriterium einen Paralleltest, so geht die Validität in die Reliabilität über: Die Reliabilität eines Tests ist dann seine Validität in Bezug auf einen Paralleltest (14).

Da bei den einzelnen Probanden das zu untersuchende Merkmal in verschiedenem Grade ausgeprägt ist, sind auch interindividuelle Unterschiede in der Testleistung zu beobachten, die in der Größe der Varianz s^2x der Rohwerte x zum Ausdruck kommen. Als Varianz wird die Summe der Abweichungsquadrate, dividiert durch die um 1 verringerte Anzahl der Beobachtungen bezeichnet (35).

3. Die Axiome der klassischen Testtheorie

Jede Messung ist mit Fehlern behaftet. Dieser Umstand findet innerhalb der klassischen Testtheorie Ausdruck in der Grundannahme, daß sich der beobachtete Meßwert X aus einem "wahren" Wert T , dem "true score", und eben jenem Meßfehler E , dem "error score" zusammensetzt (14).

$$X = T + E$$

Nach G. FISCHER (14,S.29) lauten dann die vier Axiome der klassischen Testtheorie:

"Der Erwartungswert des Fehlers ist gleich Null.

Die Korrelation zwischen Fehler und wahren Wert eines Tests ist gleich Null.

Die Korrelation zwischen Fehler im Test X_i und wahren Wert im Test X_j ist gleich Null.

Die Korrelation zwischen den Meßfehlern zweier Tests ist gleich Null."

Kurz formuliert bedeutet dies: Der durchschnittliche Meßfehler in jeder beliebigen Population ist Null, der Meßfehler ist mit dem wahren Meßwert unkorreliert, und auch die Meßfehler verschiedener Tests sind unkorreliert, sofern die Tests experimentell unabhängig voneinander vorgegeben werden. Die Zerlegung des beobachteten Meßwertes in einen wahren Wert und in den Meßfehler bietet als Nutzenanwendung unter gewissen Annahmen die Möglichkeit, den Meßfehler abzuschätzen, also Angaben über die Genauigkeit psychologischer Messungen und Testergebnisse zu machen. Solche Angaben setzen die Verfügbarkeit "äquivalenter" Messungen voraus. Je nachdem, wie einschränkend die Voraussetzungen über die Äquivalenz der Messungen sind, können mehr oder weniger präzise Angaben über den Meßfehler gemacht werden.

Eine derartige "Äquivalenz-Messung" stellt u.a. die entsprechende Untersuchung zu Parallel-Tests dar.

Zwei Messungen heißen parallel, wenn sie den gleichen Erwartungswert und die gleiche Fehlervarianz bei jeder Person aufweisen. Paralleltests erkennt man an dem gleichen Mittelwert und an der gleichen Meßwertvarianz in jeder Population. Sie messen folglich dieselbe Eigenschaft gleich gut (14).

Da diese —ideale— Voraussetzung in der Praxis nur selten erfüllt wird, verwendet man schwächere Definitionen der Äquivalenz. So sagt der Begriff der $\mathcal{T}$ -Äquivalenz, daß zwei Tests dieselbe Eigenschaft messen aber möglicherweise verschieden genau. Noch weniger strenge Annahmen liegen der essentiellen $\mathcal{T}$ -Äquivalenz zugrunde, bei der zusätzlich eine additive Konstante mit in die Berechnung eingeht.

Essentiell $\mathcal{T}$ -äquivalente Tests messen also im wesentlichen dieselbe Eigenschaft, allerdings bis auf eine additive Konstante und möglicherweise verschieden genau.

In die Formelsprache übertragen setzt sich die gemessene Varianz s^2x aus der wahren Varianz s^2t und der Fehlervarianz s^2e zusammen. Es gilt also für ein und dieselbe Stichprobe:

$$s^2x = s^2t + s^2e$$

Der Reliabilitätskoeffizient einer Messung ist dann ganz allgemein per Definitionem bestimmt durch das Verhältnis:

$$\text{Rel} = \frac{s^2t}{s^2x} ; \quad \text{da } s^2t = s^2x - s^2e \quad (14, S.27)$$

gilt:
$$\text{Rel} = 1 - \frac{s^2e}{s^2x}$$

Die Reliabilität nimmt also mit zunehmender Fehlervarianz ab. Besteht eine Messung zur Gänze aus der Fehlervarianz, dann wird die Reliabilität gleich Null, ist sie gänzlich fehlerfrei, so wird sie gleich 1. Die Formeln zeigen aber auch, daß die Reliabilität eines Tests nicht eine Eigenschaft des Tests allein ist, sondern viel eher eine Eigenschaft von Test und

Bezugspopulation zusammen. Ändert man die Varianz, indem man die Population so "homogenisiert", daß im Extremfall alle Personen den gleichen wahren Testwert aufweisen, so wird $s^2_t = 0$ und die Reliabilität verschwindet. Dies würde jetzt allerdings nicht bedeuten, daß der Test seine Meßgenauigkeit verloren hätte, sondern nur, daß Personen nicht mehr unterschieden werden können, da sie in bezug auf das durch den Test zu messende Merkmal ununterscheidbar geworden sind. Die Aussage ist also abhängig von der gewählten Population; die Reliabilität mißt die Genauigkeit des Tests in bezug auf eine gegebene Population.

4. Fehlerquellen einer Messung

In ihrer Genauigkeit beeinträchtigt werden kann eine Messung durch zwei große Klassen von Fehlerquellen: Den Mangel an Auswertungsobjektivität und den Mangel an Durchführungsobjektivität. Im ersten Falle führen eine mangelhafte Strukturiertheit der Aufgaben und eine Mehrdeutigkeit in der Lösungsbewertung zu fehlerhaften Ergebnissen, im zweiten Falle führen Veränderungen der Testdurchführungsbedingungen, Motivationschwankungen, körperliche Befindenslage und/oder Testleiterinflüsse zu Fehlern bei der Messung. Wiederholungsuntersuchungen mit Paralleltests erfordern die Berücksichtigung des potentiellen Übungstransfers. Ein Einfluß dieser Art kann ebenso eine Fehlerquelle darstellen wie Gedächtniseffekte bei Wiederholungsuntersuchungen mit dem gleichen Test.

Fehler- bzw. Störquellen erhöhen stets die unsystematische Varianz, die Fehlervarianz, die sich wiederum reliabilitätsniedrigend auswirkt.

5. Berechnung der Reliabilität

Um die Reliabilität eines Tests in praxi zu bestimmen, bedient man sich des Zusammenhanges der Ergebniswerte zweier Testuntersuchungen. Man führt also die Reliabilität auf die Korrelation zwischen zwei manifesten Variablen zurück. Dies kann auf vier verschiedene Arten geschehen (14;35).

5.1. Die Retestmethode

Die Retestreliabilität wird durch wiederholte Darbietung desselben Tests an ein und derselben Stichprobe in einem angemessenen Zeitabstand und Korrelation der Rohwertpaare aus Test und Testwiederholung ermittelt. Man erhält hier unter Umständen eine zu hohe Reliabilität, eine sog. Scheinreliabilität, weil ohne den eigentlichen Lösungsprozeß zu vollziehen, die Erinnerung den Probanden Antworten finden läßt. Die Reliabilität wird als Scheinreliabilität umso mehr in Erscheinung treten, je leichter die Testaufgaben behalten werden können, je charakteristischer und inhaltlich interessanter sie sind, je kürzer der Zeitabstand zwischen Erst- und Wiederholungstestung ist und je weniger Aufgaben der Test enthält. In der Praxis eignet sich die Reliabilitätsbestimmung über die Testwiederholungsmethode am besten dort, wo ein Test durch Lernen nicht veränderbare Fähigkeiten mißt, wie es bei reinen Schnelligkeitstesten und Fragebögen der Fall ist. Für die Reliabilitätsbestimmung von Intelligenz- und Leistungstests im Sinne des Niveaunkonzeptes sind andere Methoden besser geeignet (35).

5.2. Die Paralleltestmethode

Die Paralleltestmethode setzt das Vorhandensein zweier Parallelformen eines Tests voraus. Dies bedeutet, daß beide Formen dieselbe Eigenschaft messen, was aus der Gleichheit von Feh-

lervarianz und Mittelwert hervorgeht. Ein signifikanter Lern- oder Übungsfortschritt darf nicht bestehen. Ist dies dennoch der Fall, so liefert die Korrelation zwischen den Ergebnissen in den beiden Tests eine Schätzung der Reliabilität. Sie wird dann umso unsicherer sein, je mehr die Aufgaben einen gewissen Einmaligkeitscharakter aufweisen oder je leichter die Lösungskonzepte übertragbar sind. Bei Aufgaben vom Schnelligkeitstyp treten diese Effekte wiederum in den Hintergrund.

5.3. Die Methode der Testhalbierung

Liegt keine Parallelforn vor und besteht keine Möglichkeit der Testwiederholung, so kann die Reliabilität über die Methode der Testhalbierung ermittelt werden. Dabei wird der Test in zwei möglichst gleich lange und gleich schwierige Hälften geteilt und die Korrelation zwischen den Testteilen errechnet. Sie stellt einen Reliabilitätskoeffizienten für den halben Test dar. Zur Schätzung der Reliabilität des Gesamtttests verwendet man die bekannte SPEARMAN-BROWN-Formel (37,S.81).

5.4. Reliabilitätsschätzungen mit Hilfe des "Koeffizienten Alpha"

Reliabilitätsschätzungen können auch mit Hilfe des "Koeffizienten Alpha" durchgeführt werden. Da es in den seltensten Fällen zutrifft, daß ein Test homogen ist, also ausschließlich aus äquivalenten Items besteht, entwickelte L.J. CRONBACH (zit.14) eine Methode zur Abschätzung von Tests mit nicht äquivalenten Testteilen.

Für zwei Testteile gilt dann:

$$\begin{array}{ll} \text{gemessener Wert} & : x = x_1 + x_2 \\ \text{wahrer Wert} & : t = t_1 + t_2 \\ \text{Fehler} & : e = e_1 + e_2 \end{array}$$

nach L. GUTTMAN (zit.14) ist dann:

$$\text{Rel von } x = p^2(x,t) - 2 \left(1 - \frac{s^2 x_1 + s^2 x_2}{s^2 x} \right)$$

Für unseren Fall, d.h. für die 9 Subtests des SKT, gilt:

$$\text{Rel.} = \rho(X, T)^2 \geq \alpha = \frac{n}{n-1} \left\{ 1 - \frac{\sum \sigma^2(X_i)}{\sigma^2(X)} \right\}$$

wobei: n = Anzahl der Subtests

$s^2(x_i)$ = Varianz eines Subtests

$s^2 x$ = Varianz der Summe von n -Subtests

$\sum_{i=1}^n s^2(x_i)$ = Summe der Varianzen von n -Subtests

Der in der Ungleichung rechts verzeichnete Ausdruck wird als "Koeffizient Alpha" bezeichnet und stellt eine untere Schranke der Reliabilität dar. Das Gleichheitszeichen in der oben dargestellten Ungleichung gilt nur dann, wenn folgende Bedingungen zutreffen:

Der "Koeffizient Alpha" ist gleich der Reliabilität, wenn die Testteile parallel sind.

Der "Koeffizient Alpha" ist genau dann gleich der Reliabilität, wenn die Testteile essentiell $\mathcal{V}$ -äquivalent sind.

In der Praxis können Tests häufig so geteilt werden (oder es liegen Parallelformen vor), daß die Testteile dieselbe Eigenschaft bis auf eine verschiedene Meßgenauigkeit und eine additive Konstante messen. Der "Koeffizient Alpha" liefert daher eine Abschätzung der Reliabilität, die umso genauer ist, je mehr die getroffenen Annahmen an die Realität heranreichen.

Im Rahmen der Reliabilitätsbestimmung wird schließlich noch der Begriff der "strengen Parallelität" von dem der "nominellen Parallelität" unterschieden. Hier postuliert man zu einem Test X eine ganze Familie (ein "Universum") von nominell parallelen Tests mit testspezifischen "true-scores" sowie "error-scores" und definiert die mittlere Testleistung einer Versuchsperson in einem zufällig aus dem Testuniversum ausgewählten Test als den "universe-true-score" bzw. "universe-error-score". Legt man nun eine Stichprobe von $k \geq 2$ nominell parallelen Tests einer Stichprobe von n Probanden vor, so kann man die beobachteten Meßwerte einer Varianzanalyse zuführen (14).

Dort, für den Spezialfall "k streng paralleler" Tests, gilt: Die Reliabilität des Tests i ist gleich der wahren Varianz dividiert durch die Testvarianz:

$$Rel = p^2(x, t) = \frac{s^2 t_i}{s^2 t_i + s^2 e_i} ;$$

Eine Schätzung erhält man, indem die Varianzen in Zähler und Nenner durch erwartungstreue Varianzschätzungen nach folgender Tafel ersetzt werden:

Varianzkomponenten	Abweichungsquadratsummen	df	Erwartungswert des mittleren Abweichungs- quadrates
Zwischen den Personen	$AQZ = k \sum_{v=1}^n (\bar{x}_{v.} - \bar{x}_{..})^2$	$n - 1$	$\sigma_E^2 + k\sigma_T^2$
Innerhalb (Replikationen)	$AQI = \sum_{v=1}^n \sum_{i=1}^k (x_{vi} - \bar{x}_{v.})^2$	$n(k - 1)$	σ_E^2
Total	$AQT = \sum_{v=1}^n \sum_{i=1}^k (x_{vi} - \bar{x}_{..})^2$	$nk - 1$	

Abb. IV.1. Tafel der Varianzanalyse für "k streng parallele" Tests

Formeln für die praktische Berechnung:

$$AQT = \sum_v \sum_i x_{vi}^2 - \frac{1}{nk} \left(\sum_v \sum_i x_{vi} \right)^2$$

$$AQZ = \frac{1}{k} \sum_v \left(\sum_i x_{vi} \right)^2 - \frac{1}{nk} \left(\sum_v \sum_i x_{vi} \right)^2$$

$$AQI = AQT - AQZ$$

Abb. IV.1

Tafel der Varianzanalyse für "k streng parallele" Tests
(Fortsetzung)

Man erhält die Formel:

$$\hat{\rho}_{XT}^2 = \frac{AQZ - \frac{1}{k-1} \frac{n-1}{n} AQI}{AQZ + \frac{n-1}{n} AQI}$$

Für den Spezialfall "k nominell paralleler" Tests gilt:
Die spezifische Reliabilität des i-ten Tests ist:

$$Rel = p^2(x, t) = 1 - \frac{s^2 e_i}{s^2 X_i} ;$$

Eine Varianzanalyse wird wiederum mit der Tafel auf der folgenden Seite durchgeführt:

Varianzkomponenten	Abweichungsquadratsummen	df	Erwartungswerte
Zwischen Personen	$AQZ_P = k \sum_{v=1}^n (\bar{x}_v - \bar{x}_{..})^2$ $= \frac{1}{k} \sum_v \left(\sum_i x_{vi} \right)^2 - \frac{1}{nk} \left(\sum_v \sum_i x_{vi} \right)^2$	$n - 1$	$\sigma^2(E_{ee}) + (1 - \frac{k}{K})\sigma_a^2 + k\sigma_{\xi}^2$
Zwischen Tests	$AQZ_T = n \sum_{i=1}^k (\bar{x}_{.i} - \bar{x}_{..})^2$ $= \frac{1}{n} \sum_i \left(\sum_v x_{vi} \right)^2 - \frac{1}{nk} \left(\sum_v \sum_i x_{vi} \right)^2$	$k - 1$	$\sigma^2(E_{ee}) + \sigma_a^2 + n\sigma_{\pi}^2$
Restvarianz	$AQR = \sum_v \sum_i (x_{vi} - \bar{x}_v - \bar{x}_{.i} + \bar{x}_{..})^2$	$(n - 1)(k - 1)$	$\sigma^2(E_{ee}) + \sigma_a^2$
Total	$AQT = \sum_v \sum_i (x_{vi} - \bar{x}_{..})^2$ $= \sum_v \sum_i x_{vi}^2 - \frac{1}{nk} \left(\sum_v \sum_i x_{vi} \right)^2$	$nk - 1$	

Abb. IV.2

Tafel der Varianzanalyse für "k nominell parallele" Tests

Als Formel für die Schätzung der Reliabilität erhält man:

$$\hat{\rho}_{XT}^2 = 1 - \frac{\frac{1}{(k-1)(n-1)} AQR}{\frac{1}{n-1} \sum_{v=1}^n (x_{vi} - \bar{x}_{.i})^2}$$

V Testdurchführung und Befunde (Rohdaten)

1. Durchführung der Untersuchung

Der SKT der Formen A bis E (A wurde zweimal vorgelegt) ist im Zeitraum August bis Oktober 1979 an 120 Patienten mehrerer Kliniken der Universität Erlangen abgenommen worden. Die Kranken waren Patienten der chirurgischen, der neurologischen und der neurochirurgischen Kliniken. Die Auswahl der Probanden wurde nach folgenden Kriterien vorgenommen: Damit Wiederholungstestungen garantiert waren, sollte die voraussichtliche Verweildauer auf den entsprechenden Stationen mindestens eine Woche betragen. Das Alter der Patienten hatte gemäß den bei der Normierung vorgegebenen Altersgruppen zwischen 17 und 54 Jahren zu sein. Der Intelligenzquotient mußte in einem Bereich zwischen 91 und 112 (jeweils einschließlich) liegen (vgl. Abb. III.2, S.16).

Weiterhin war zu beachten, daß im Rahmen der Retestreliabilitätsbestimmung die Ergebnisse einer Wiederholungsmessung ja nur dann verwertbar sind, wenn man annehmen kann, daß sich das zu messende Merkmal in der Zwischenzeit nicht verändert hat. Dieser Forderung nach einer einheitlichen Testgruppe vom Typus "relative (Leistungs-) Merkmalskonstanz" wurde durch Ausschluß einer psychiatrischen Anamnese und einer aktuellen psychischen Erkrankung Rechnung getragen. Die Testuntersuchung ist in den meisten Fällen am Krankbett vorgenommen worden. Wenn es sich ermöglichen ließ, wichen wir in eine "neutrale" Umgebung aus (Untersuchungszimmer oder Aufenthaltsräume). Der Untersuchungstermin

wurde gern so gewählt, daß keine Störung durch Essensausgabe, Besuchszeiten oder anderweitige Untersuchungen stattfand.

Zunächst war es erforderlich, mit Hilfe des MWT-B das allgemeine Intelligenzniveau grob abzuschätzen. Nach dem bereits genannten Auswahl Gesichtspunkt fanden ausschließlich Personen mit einer erreichten Rohpunktzahl zwischen 20 und 31 (jeweils einschließlich) Aufnahme in die Testgruppe. Es handelt sich bei diesen Patienten also um "durchschnittlich intelligente" Personen mit einem IQ zwischen 91 und 112 (jeweils einschließlich, vgl. Abb. III.2, S.16).

Um eine möglichst standardisierte Testsituation zu gewährleisten, achteten wir auf eine zügige Durchführung der Untersuchung ohne größere Pausen. Die Testdauer betrug in der Regel ca. 50 Minuten.

Um etwaige Positionseffekte bei der Vorlage der Paralleltests kontrollieren zu können, wurde die Reihenfolge der Testformen A bis E nach einem Rotationsschema variiert. Bei 24 Probanden (einem Fünftel der Gesamtgruppe) wurde mit Testform A begonnen, danach jeweils fortlaufend für dieselbe Personenzahl B bis E. Auf diese Weise befanden sich alle Testformen mit gleicher Häufigkeit am selben Platz der Testreihenfolge. Nach einer Woche wurde eine Testwiederholung durchgeführt. Hier beschränkten wir uns auf die Testform A.

Damit eine möglichst gute Motivation zu gewährleisten war, informierten wir in groben Zügen über Sinn und Zweck der Untersuchung. Danach baten wir die Patienten um Zustimmung und Mitarbeit. Diese Maßnahmen in Verbindung mit dem spielerisch gestalteten Material des SKT dürften für die geringe Zurückweisungsquote von ca. 3% verantwortlich sein (von 124 befragten Kranken hatten nur 4 die Mitarbeit abgelehnt).

2. Beschreibung der Untersuchungsstichprobe

Die 120 getesteten Patienten sind in eine Gesamtgruppe und in eine zweite Gruppe aufgeteilt worden, bei deren Auswahl von folgender Überlegung ausgegangen wurde: Da Testform A das in der Praxis am häufigsten angewandte und statistisch am besten abgesicherte Einzelverfahren darstellt, wurden alle Probanden, die in diesem Test weniger als 5 Punkte erzielt hatten (klinische Klassifizierung: kein Durchgangssyndrom nachweisbar) einer zweiten Gruppe zugeordnet. In diese Kategorie gelangten insgesamt 76 Personen.

Das Durchschnittsalter der Patienten und die Geschlechtsverteilung für die Gruppen $n = 120$ und $n = 76$ ist der folgenden Tabelle zu entnehmen.

Tabelle V.1. Durchschnittsalter (mit Standardabweichung) und Geschlechtsverteilung der untersuchten Patienten

	Durchschnittsalter (M)		SD	
	absolut	in Prozent	absolut	in Prozent
$n = 120$	64	53,3	56	46,7
$n = 76$	35	46,1	41	53,9

Unter Berücksichtigung der Grunderkrankung und des klinischen Bildes wurde versucht, eine grobe Einteilung der Probanden in Patienten mit fraglicher, wahrscheinlicher oder

ausgeschlossener "Beteiligung zerebraler Bereiche" zu geben.

Tabelle V.2

Die Häufigkeitsverteilung der zerebralen Beteiligung beider Gruppen

	n = 120	n = 76
wahrscheinliche zerebrale Beteiligung	9	3
fragliche " "	102	71
keine " "	9	2
Gesamt	120	76

Über die Verteilung der Patienten auf die einzelnen Kliniken gibt die nachfolgende Tabelle Auskunft.

Tabelle V.3

Die Herkunft der Patienten aus den einzelnen Kliniken

	n = 120	n = 76
Neurochirurgische Klinik	14	6
Neurologische Klinik	33	15
Psychiatrische Klinik	13	8
Chirurgische Klinik	60	47
Gesamt	120	76

Die grobe Schätzung des Intelligenzniveaus wurde mit Hilfe des MWT-B durchgeführt. Wie bereits erwähnt, fanden ausschließlich Probanden mit einer Rohpunktzahl zwischen 20 und 31 (jeweils einschließlich) Aufnahme in das Testkollektiv. Die Mittelwerte dieser im Durchschnittsbereich

liegenden Leistungen zeigt Tabelle V.4.

Tabelle V.4. Die Mittelwerte der Rohpunkte mit Standardabweichung und zugehörigem Mittelwert der Intelligenzquotienten für beide Gruppen.

n (Gruppe 1 u. 2)	Rohpunkte (M)	SD	IQ (M)
n = 120	26,88	3,359	100
n = 76	27,62	3,149	101

Die Rohwerte aus den Subtests der Testformen A bis E einschließlich der Testwiederholung von A für die Gruppen n = 120 und n = 76 sind in Tabelle V.5 auf den nächsten Seiten dargestellt.

Die in den einzelnen Subtests erreichten Rohwerte wurden entsprechend den im Manual zum SKT (10) verzeichneten Normwerten in Wertpunkte umgerechnet (vgl. Tabelle V.6 auf den Seiten 44 und 45).

Die Häufigkeitsverteilung der im Gesamttest (Formen A bis E) einschließlich der wiederholten Untersuchung mit Form A erzielten Wertpunkte bietet Tabelle V.7 ab Seite 46.

Tabelle V.5a

Die Mittelwerte der Rohpunkte sämtlicher Subtests einschließlich Standardabweichung der Gruppe n = 76

		MEANS	STD DEV	CASES
1.	RCHA1	11.59211	2.31043	76.0
2.	RCHA2	5.13158	1.42681	76.0
3.	RCHA3	6.46053	1.14823	76.0
4.	RCHA4	15.86842	3.48077	76.0
5.	RCHA5	13.35526	2.43148	76.0
6.	RCHA6	17.43421	3.60494	76.0
7.	RCHA7	19.73684	3.85528	76.0
8.	RCHA8	4.80263	1.58341	76.0
9.	RCHA9	0.61842	0.72970	76.0
1.	RCHB1	11.72368	2.59023	76.0
2.	RCHB2	5.59211	1.58474	76.0
3.	RCHB3	6.44737	1.07573	76.0
4.	RCHB4	16.48684	3.55713	76.0
5.	RCHB5	13.85526	2.60105	76.0
6.	RCHB6	17.92105	3.24042	76.0
7.	RCHB7	20.77632	3.54626	76.0
8.	RCHB8	5.19737	2.34390	76.0
9.	RCHB9	0.60526	0.83393	76.0
1.	RCHC1	11.21053	2.93174	76.0
2.	RCHC2	5.57895	1.46299	76.0
3.	RCHC3	6.53947	1.28001	76.0
4.	RCHC4	15.88158	3.35844	76.0
5.	RCHC5	13.43421	2.86280	76.0
6.	RCHC6	16.31579	3.81300	76.0
7.	RCHC7	19.03947	3.83472	76.0
8.	RCHC8	5.00000	2.00000	76.0
9.	RCHC9	0.40789	0.59279	76.0
1.	RCHD1	13.75000	3.17543	76.0
2.	RCHD2	5.64474	1.42083	76.0
3.	RCHD3	6.35526	1.13964	76.0
4.	RCHD4	15.86842	3.48077	76.0
5.	RCHD5	13.80263	2.81908	76.0
6.	RCHD6	17.02632	3.65504	76.0
7.	RCHD7	20.14474	3.50125	76.0
8.	RCHD8	4.97368	2.25077	76.0
9.	RCHD9	0.42105	0.63798	76.0
1.	RCHE1	13.10526	3.55370	76.0
2.	RCHE2	5.55263	1.22632	76.0
3.	RCHE3	6.07895	0.93471	76.0
4.	RCHE4	15.21053	3.26728	76.0
5.	RCHE5	12.51316	2.20601	76.0
6.	RCHE6	20.69737	3.97666	76.0
7.	RCHE7	18.09211	3.12912	76.0
8.	RCHE8	4.72368	2.28968	76.0
9.	RCHE9	0.28947	0.48486	76.0
1.	RCHAW1	10.72368	2.65128	76.0
2.	RCHAW2	3.78684	1.23821	76.0
3.	RCHAW3	6.44737	1.13601	76.0
4.	RCHAW4	14.93421	3.37574	76.0
5.	RCHAW5	13.59211	3.22150	76.0
6.	RCHAW6	17.26316	3.96776	76.0
7.	RCHAW7	18.63158	4.45972	76.0
8.	RCHAW8	3.46053	1.62842	76.0
9.	RCHAW9	0.31579	0.63688	76.0

Tabelle V.5b

Die Mittelwerte der Rohpunkte sämtlicher Subtests einschließlich Standardabweichung der Gruppe n = 120

		MEANS	STD DEV	CASES
1.	RDHA1	13.10833	3.70509	120.0
2.	RDHA2	5.75000	1.59963	120.0
3.	RDHA3	6.80833	1.45114	120.0
4.	RDHA4	18.53333	5.66863	120.0
5.	RDHA5	15.18333	4.31332	120.0
6.	RDHA6	18.97500	4.67869	120.0
7.	RDHA7	21.08333	6.17393	120.0
8.	RDHA8	5.57500	1.93448	120.0
9.	RDHA9	0.98333	1.53931	120.0
10.	RDHB1	12.91667	3.61273	120.0
11.	RDHB2	5.94167	1.61034	120.0
12.	RDHB3	6.85833	1.45692	120.0
13.	RDHB4	19.05000	5.47546	120.0
14.	RDHB5	15.57500	4.12140	120.0
15.	RDHB6	19.35000	4.57918	120.0
16.	RDHB7	23.12500	5.95157	120.0
17.	RDHB8	5.76667	2.34694	120.0
18.	RDHB9	0.96667	1.47206	120.0
19.	RDHC1	12.47500	4.39817	120.0
20.	RDHC2	5.87500	1.59074	120.0
21.	RDHC3	6.95833	1.80287	120.0
22.	RDHC4	18.89167	6.12585	120.0
23.	RDHC5	15.05000	3.77486	120.0
24.	RDHC6	17.69167	5.18043	120.0
25.	RDHC7	21.41667	6.29870	120.0
26.	RDHC8	5.60000	2.12785	120.0
27.	RDHC9	0.70833	1.24614	120.0
28.	RDHD1	15.29333	4.63850	120.0
29.	RDHD2	5.95000	1.42516	120.0
30.	RDHD3	6.73333	1.49921	120.0
31.	RDHD4	18.44167	5.90031	120.0
32.	RDHD5	15.69167	4.68826	120.0
33.	RDHD6	18.55000	4.95161	120.0
34.	RDHD7	22.68333	6.44561	120.0
35.	RDHD8	5.49167	2.24917	120.0
36.	RDHD9	0.72500	1.03682	120.0
37.	RDHE1	14.52500	4.83651	120.0
38.	RDHE2	6.01667	1.46088	120.0
39.	RDHE3	6.40833	1.25354	120.0
40.	RDHE4	17.51667	5.58808	120.0
41.	RDHE5	14.18333	3.85228	120.0
42.	RDHE6	22.01667	5.54278	120.0
43.	RDHE7	20.34167	6.08966	120.0
44.	RDHE8	5.55000	2.35451	120.0
45.	RDHE9	0.69167	1.08307	120.0
46.	RDHAW1	11.65000	3.53054	120.0
47.	RDHAW2	4.30000	1.35721	120.0
48.	RDHAW3	6.70000	1.36339	120.0
49.	RDHAW4	17.40833	5.67539	120.0
50.	RDHAW5	15.05833	4.33725	120.0
51.	RDHAW6	18.38333	4.65839	120.0
52.	RDHAW7	20.41667	5.59949	120.0
53.	RDHAW8	3.81667	1.73923	120.0
54.	RDHAW9	0.46667	0.75519	120.0

Tabelle V.6a

Umrechnung der Rohpunkte in Wertpunkte

Alter: 17 bis 44 Jahre

IQ (MWT-B): 90 bis 110

Wert- punkte	Subtest	I sec	II Fehler	III sec	IV sec	V sec	VI sec	VII sec	VIII Fehler	IX Fehler
0		bis 13	bis 5	bis 6	bis 18	bis 16	bis 17	bis 24	bis 4	1
1		14 - 17	6 - 7	7 - 9	19 - 27	17 - 25	18 - 26	25 - 35	5 - 7	2 - 3
2		18 - 22	8 - 9	10 - 13	28 - 35	26 - 37	27 - 51	37 - 55	8 - 9	4 - 7
3		23 - 60	10 - 12	14 - 60	36 - 60	38 - 60	52 - 60	56 - 60	10 - 12	8 - 10

Tabelle V.6b

Umrechnung der Rohpunkte in Wertpunkte

Alter: 45 bis 54 Jahre

IQ (MWT-B): 90 bis 110

Wert- punkte	Subtest	I sec	II Fehler	III sec	IV sec	V sec	VI sec	VII sec	VIII Fehler	IX Fehler
0		bis 13	bis 5	bis 7	bis 23	bis 22	bis 19	bis 27	bis 5	bis 1
1		14 - 22	6 - 7	8 - 13	24 - 30	23 - 28	20 - 26	28 - 42	6 - 7	2 - 3
2		23 - 40	8 - 9	14 - 20	31 - 45	29 - 46	27 - 39	43 - 55	8 - 9	4 - 5
3		41 - 60	10 - 12	21 - 60	46 - 60	47 - 60	40 - 60	56 - 60	10 - 12	6 - 12

Tabelle V.7a. Die Häufigkeitsverteilung (hier: ABSOLUTE FREQ u. RELATIVE FREQ) der im Gesamttest A u. B für n = 120 erzielten Wertpunkte (CODE)

GESA					
CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	10	8,3	8,3	8,3
	1.	10	8,3	8,3	16,7
	2.	24	20,0	20,0	36,7
	3.	18	15,0	15,0	51,7
	4.	14	11,7	11,7	63,3
	5.	12	10,0	10,0	73,3
	6.	4	3,3	3,3	76,7
	7.	10	8,3	8,3	85,0
	8.	5	4,2	4,2	89,2
	9.	5	4,2	4,2	93,3
	10.	4	3,3	3,3	96,7
	11.	1	0,8	0,8	97,5
	12.	1	0,8	0,8	98,3
	15.	1	0,8	0,8	99,2
	18.	1	0,8	0,8	100,0
	TOTAL	120	100,0	100,0	

GESB					
CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	6	5,0	5,0	5,0
	1.	11	9,2	9,2	14,2
	2.	14	11,7	11,7	25,8
	3.	21	17,5	17,5	43,3
	4.	19	15,8	15,8	59,2
	5.	17	14,2	14,2	73,3
	6.	7	5,8	5,8	79,2
	7.	7	5,8	5,8	85,0
	8.	6	5,0	5,0	90,0
	9.	5	4,2	4,2	94,2
	10.	4	3,3	3,3	97,5
	13.	1	0,8	0,8	98,3
	14.	1	0,8	0,8	99,2
	18.	1	0,8	0,8	100,0
	TOTAL	120	100,0	100,0	

Tabelle V.7b. Die Häufigkeitsverteilung (hier: ABSOLUTE FREQ u. RELATIVE FREQ) der im Gesamttest C u. D für n = 120 erzielten Wertpunkte (CODE)

GES C

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	8	6,7	6,7	6,7
	1.	14	11,7	11,7	18,3
	2.	22	18,3	18,3	36,7
	3.	31	25,8	25,8	62,5
	4.	8	6,7	6,7	69,2
	5.	6	5,0	5,0	74,2
	6.	9	7,5	7,5	81,7
	7.	7	5,8	5,8	87,5
	8.	4	3,3	3,3	90,8
	9.	3	2,5	2,5	93,3
	10.	1	0,8	0,8	94,2
	11.	4	3,3	3,3	97,5
	13.	3	2,5	2,5	100,0
	TOTAL	120	100,0	100,0	

GES D

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	3	2,5	2,5	2,5
	1.	15	12,5	12,5	15,0
	2.	19	15,8	15,8	30,8
	3.	21	17,5	17,5	48,3
	4.	12	10,0	10,0	58,3
	5.	14	11,7	11,7	70,0
	6.	12	10,0	10,0	80,0
	7.	4	3,3	3,3	83,3
	8.	6	5,0	5,0	88,3
	9.	4	3,3	3,3	91,7
	10.	3	2,5	2,5	94,2
	12.	4	3,3	3,3	97,5
	13.	1	0,8	0,8	98,3
	14.	1	0,8	0,8	99,2
	15.	1	0,8	0,8	100,0
	TOTAL	120	100,0	100,0	

Tabelle V.7c. Die Häufigkeitsverteilung (hier: ABSOLUTE FREQ u. RELATIVE FREQ) der im Gesamttest E u. AW für n = 120 erzielten Wertpunkte (CODE)

GESE

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	9	7,5	7,5	7,5
	1.	8	6,7	6,7	14,2
	2.	21	17,5	17,5	31,7
	3.	19	15,8	15,8	47,5
	4.	19	15,8	15,8	63,3
	5.	7	5,8	5,8	69,2
	6.	11	9,2	9,2	78,3
	7.	4	3,3	3,3	81,7
	8.	8	6,7	6,7	88,3
	9.	4	3,3	3,3	91,7
	10.	2	1,7	1,7	93,3
	11.	4	3,3	3,3	96,7
	12.	1	0,8	0,8	97,5
	13.	2	1,7	1,7	99,2
	15.	1	0,8	0,8	100,0
TOTAL		120	100,0	100,0	

GESAMT

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	24	20,0	20,0	20,0
	1.	31	25,8	25,8	45,8
	2.	24	20,0	20,0	65,8
	3.	15	12,5	12,5	78,3
	4.	7	5,8	5,8	84,2
	5.	8	6,7	6,7	90,8
	6.	3	2,5	2,5	93,3
	8.	2	1,7	1,7	95,0
	9.	2	1,7	1,7	96,7
	10.	1	0,8	0,8	97,5
	11.	1	0,8	0,8	98,3
	12.	1	0,8	0,8	99,2
	14.	1	0,8	0,8	100,0
TOTAL		120	100,0	100,0	

Tabelle V.7d. Die Häufigkeitsverteilung (hier: ABSOLUTE FREQ u. RELATIVE FREQ) der im Gesamttest A, B u. C für n = 76 erzielten Wertpunkte (CODE)

GESA

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	10	13.2	13.2	13.2
	1.	10	13.2	13.2	26.3
	2.	24	31.6	31.6	57.9
	3.	18	23.7	23.7	81.6
	4.	14	18.4	18.4	100.0
	TOTAL	76	100.0	100.0	

GESB

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	6	7.9	7.9	7.9
	1.	11	14.5	14.5	22.4
	2.	14	18.4	18.4	40.8
	3.	19	25.0	25.0	65.8
	4.	16	21.1	21.1	86.8
	5.	8	10.5	10.5	97.4
	6.	1	1.3	1.3	98.7
	9.	1	1.3	1.3	100.0
	TOTAL	76	100.0	100.0	

GESC

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	8	10.5	10.5	10.5
	1.	13	17.1	17.1	27.6
	2.	19	25.0	25.0	52.6
	3.	27	35.5	35.5	88.2
	4.	3	3.9	3.9	92.1
	5.	3	3.9	3.9	96.1
	6.	2	2.6	2.6	98.7
	8.	1	1.3	1.3	100.0
	TOTAL	76	100.0	100.0	

Tabelle V.7e. Die Häufigkeitsverteilung (hier: ABSOLUTE FREQ u. RELATIVE FREQ) der im Gesamttest D, E u. AW für n = 76 erzielten Wertpunkte (CODE)

GESD

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	3	3.9	3.9	3.9
	1.	15	19.7	19.7	23.7
	2.	18	23.7	23.7	47.4
	3.	18	23.7	23.7	71.1
	4.	9	11.8	11.8	82.9
	5.	9	11.8	11.8	94.7
	6.	4	5.3	5.3	100.0
TOTAL		76	100.0	100.0	

GESE

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	9	11.8	11.8	11.8
	1.	8	10.5	10.5	22.4
	2.	20	26.3	26.3	48.7
	3.	18	23.7	23.7	72.4
	4.	13	17.1	17.1	89.5
	5.	4	5.3	5.3	94.7
	6.	3	3.9	3.9	98.7
	7.	1	1.3	1.3	100.0
TOTAL		76	100.0	100.0	

GESAW

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	23	30.3	30.3	30.3
	1.	25	32.9	32.9	63.2
	2.	17	22.4	22.4	85.5
	3.	6	7.9	7.9	93.4
	4.	3	3.9	3.9	97.4
	5.	1	1.3	1.3	98.7
	6.	1	1.3	1.3	100.0
TOTAL		76	100.0	100.0	

VI Ergebnisse

Die statistischen Auswertungen und Berechnungen der Rohdaten sind an der TR 440 des Regionalen Rechenzentrums Erlangen (RRZE) durchgeführt worden. Es wurden dabei aus dem SPSS-7 (18,19) folgende Programme verwendet:

Condescriptive,
Pearson Corr.,
Oneway,
Reliability.

An dieser Stelle sei nochmals Herrn Dipl.Psych. K. ARNOLD für die statistische Beratung und die Arbeiten im Rechenzentrum gedankt.

1. Reliabilitätsberechnungen

Die vorgegebenen Rohdaten wurden einer Varianzanalyse zu-geführt.

Vorgehensweise: Annahme strenger Parallelität

Die Retestreliabilität kann nach der folgenden Formel berechnet werden (14):

$$\hat{\rho}_{XT}^2 = \frac{AQZ - \frac{1}{k-1} \frac{n-1}{n} AQI}{AQZ + \frac{n-1}{n} AQI}$$

Für die Gruppen $n = 120$ und $n = 76$ ergeben sich dann die in Tab. VI.1 dargestellten Reliabilitätskoeffizienten.

Tabelle VI.1

Die Retestreliabilitätskoeffizienten (XT) der einzelnen Subtests unter der Annahme strenger Parallelität für die Gruppen $n = 120$ und $n = 76$ (AQI bzw. AQZ = Varianzkomponenten "zwischen" bzw. "innerhalb" den Personen, vgl. S.34,35)

$n = 120$

Subtest	AQI	AQZ	XT
1 A_1/AW_1	780,500	2463,996	0,522
2 A_2/AW_2	318,000	331,850	0,026
3 A_3/AW_3	81,500	390,996	0,656
4 A_4/AW_4	1317,500	6415,296	0,662
5 A_5/AW_5	797,500	3655,996	0,644
6 A_6/AW_6	472,500	4735,796	0,820
7 A_7/AW_7	921,000	7493,400	0,783
8 A_8/AW_8	444,500	546,296	0,107
9 A_9/AW_9	137,000	228,850	0,255

$n = 76$

Subtest	AQI	AQZ	XT
1 A_1/AW_1	262,000	694,211	0,457
2 A_2/AW_2	160,500	156,967	-0,005
3 A_3/AW_3	25,500	170,178	0,742
4 A_4/AW_4	540,500	1256,020	0,404
5 A_5/AW_5	395,000	828,895	0,360
6 A_6/AW_6	254,500	1902,020	0,767
7 A_7/AW_7	373,000	2279,842	0,722
8 A_8/AW_8	205,000	250,368	0,106
9 A_9/AW_9	30,500	43,336	0,180

Für beide Gruppen weist jeweils Subtest 2 den geringsten (für $n = 76$ erscheint sogar ein neg. Wert, siehe Diskussion) und Subtest 6 den höchsten Retestreliabilitätskoeffizienten auf.

Vorgehensweise: Annahme nomineller Parallelität

Die Retestreliabilität läßt sich auch nach folgender Formel (14) berechnen:

$$\hat{\rho}_{XT}^2 = 1 - \frac{\frac{1}{(k-1)(n-1)} \text{AQR}}{\frac{1}{n-1} \sum_{v=1}^n (x_{vi} - \bar{x}_{.i})^2}$$

Für die Gruppen $n = 120$ und $n = 76$ ergeben sich die in Tabelle VI.2 aufgezeichneten Reliabilitätskoeffizienten. XT_1 stellt dabei die Reliabilitätskoeffizienten für die Testform A, XT_2 die Reliabilitätskoeffizienten für die Testwiederholung AW dar.

Tabelle VI.2

Die Retestreliabilitätskoeffizienten (XT_1 für Testform A, XT_2 für Testwiederholung AW) der einzelnen Subtests unter der Annahme nomineller Parallelität für die Gruppen $n = 120$ und $n = 76$ (einschließlich zugehöriger Varianzen, vgl. S.36)

$n = 120$

Subtest	AQR	Varianzen		XT_1	XT_2
1 A_1/AW_1	652,896	13,728	12,465	0,600	0,560
2 A_2/AW_2	191,850	2,559	1,842	0,370	0,125
3 A_3/AW_3	80,796	2,106	1,859	0,678	0,635
4 A_4/AW_4	1241,563	32,133	32,210	0,675	0,676
5 A_5/AW_5	796,563	18,605	18,812	0,640	0,644
6 A_6/AW_6	451,496	21,890	21,701	0,827	0,825
7 A_7/AW_7	773,733	38,117	31,354	0,829	0,793
8 A_8/AW_8	258,996	3,740	3,025	0,418	0,281
9 A_9/AW_9	120,983	2,370	0,570	0,571	0,783

$n = 76$

Subtest	AQR	Varianzen		XT_1	XT_2
1 A_1/AW_1	233,342	5,338	7,029	0,417	0,557
2 A_2/AW_2	110,704	2,036	1,533	0,275	0,037
3 A_3/AW_3	25,493	1,318	1,291	0,742	0,737
4 A_4/AW_4	507,336	12,116	11,396	0,442	0,406
5 A_5/AW_5	392,868	10,778	10,378	0,514	0,495
6 A_6/AW_6	253,388	12,996	15,743	0,740	0,785
7 A_7/AW_7	326,579	14,863	19,889	0,707	0,781
8 A_8/AW_8	136,553	2,507	2,652	0,274	0,313
9 A_9/AW_9	27,020	0,533	0,406	0,323	0,112

Für die Gruppe $n = 120$ weisen Subtest 2 die niedrigsten und Subtests 6 u. 7 die höchsten Werte auf. Dasselbe gilt für die Gruppe $n = 76$ mit Ausnahme von Subtest 3, der ebenfalls hohe Werte zeigt.

Vorgehensweise: Schätzung nach L.J. CRONBACH

Die Abschätzung der Reliabilität für die Gesamttestform A bzw. AW (= das Gesamtergebnis von Form A bzw. AW über alle 9 Einzelaufgaben hinweg) ist mit Hilfe der Methode nach L.J. CRONBACH (vgl.14) vorgenommen worden. Im Gegensatz zu der Berechnung des "non standardized item Alpha" werden bei der Berechnung des "standardized item Alpha" die Skaleneinflüsse, die durch die unterschiedlichen Skalierungen der Subtests entstehen (bei Gedächtnisaufgaben 0 bis 12 Punkte, bei Schnelligkeitsaufgaben 0 bis 60 Sekunden), berücksichtigt. Den "exakteren" Wert liefert also das "standardized item Alpha".

Tabelle VI.3

Das "non standardized item Alpha" und "standardized item Alpha" nach L.J. CRONBACH der Testform A und Testwiederholung AW für die Gruppen n = 120 und n = 76

n = 120		
	non standardized item Alpha	standardized item Alpha
Testform A	0,81566	0,84017
Testform AW	0,79823	0,82421
n = 76		
	non standardized item Alpha	standardized item Alpha
Testform A	0,58698	0,60256
Testform AW	0,68771	0,68111

Das standardized item Alpha liefert also -bis auf Testform AW der Gruppe n = 76- die deutlich besseren Werte.

2. Korrelationen

Unter bestimmten Bedingungen (hier: "true-score" = "error-score", vgl. S.34) können die Korrelationen der entsprechenden Subtests als Reliabilitäten interpretiert werden.

Die folgende Tabelle zeigt die Korrelationskoeffizienten zwischen Testform A und Testwiederholung AW für die Subtests 1 bis 9.

Tabelle VI.4

Korrelationskoeffizienten zwischen Testform A und Testwiederholung AW für alle 9 Subtests

	Subtest	Korrelation
1	A_1/AW_1	0,5817
2	A_2/AW_2	0,2709
3	A_3/AW_3	0,6588
4	A_4/AW_4	0,6757
5	A_5/AW_5	0,6422
6	A_6/AW_6	0,8259
7	A_7/AW_7	0,8167
8	A_8/AW_8	0,3588
9	A_9/AW_9	0,3899

Den geringsten Retestrelabilitätswert weist hier Subtest 2 mit 0,27, den höchsten Subtest 6 mit 0,83 auf.

3. Einflüsse durch Alter und Intelligenz

Auf den folgenden zwei Seiten sind in einer Tabelle die Korrelationskoeffizienten zwischen den Variablen "Alter" sowie "Intelligenz" einerseits und sämtlichen (item = Einzelaufgaben-) Rohdaten aller Testformen einschließlich Testwiederholung bei Form A andererseits dargestellt (Tabellen VI.5a und VI.5b).

Als Kriterium für eine deutliche Signifikanz wurde ein Niveau von ≤ 0.01 angenommen. Die Abbildungen VI.1a und VI.1b zeigen innerhalb der jeweiligen Darstellung in Tabellenform die entsprechenden signifikanten Korrelationen durch ein "x" herausgehoben auf.

Alter:		Testform					
		A	B	C	D	E	AW
Subtest	1					x	
"	2						
"	3						
"	4						
"	5						x
"	6						x
"	7			x			x
"	8				x	x	
"	9						

Abb. VI.1a

Intelligenz:		Testform					
		A	B	C	D	E	AW
Subtest	1	x	x		x	x	
"	2						
"	3						
"	4	x	x			x	x
"	5	x	x	x	x		x
"	6						
"	7		x	x	x	x	x
"	8						
"	9						

Abb. VI.1b

Die Abb. VI.1 a und b zeigen die signifikanten Korrelationen zwischen den Rohdaten sämtlicher Subtests und den Variablen Alter sowie Intelligenz

Tabelle VI.5a. Korrelationen zwischen den Variablen Alter sowie Intelligenz und den Rohdaten der Testformen A bis C

	PEARSON CORRELATION			COEFFICIENTS		
	RHA1	RHA2	RHA3	RHA4	RHA5	RHA6
AGE	0.2023 (120) S=0.027	0.2053 (120) S=0.024	0.1101 (120) S=0.231	0.1117 (120) S=0.224	0.1058 (120) S=0.250	0.1460 (120) S=0.112
MWT	-0.2360 (120) S=0.009	-0.2182 (120) S=0.017	-0.0477 (120) S=0.605	-0.3330 (120) S=0.001	-0.2787 (120) S=0.002	-0.1841 (120) S=0.044
	PEARSON CORRELATION			COEFFICIENTS		
	RHB1	RHB2	RHB3	RHB4	RHB5	RHB6
AGE	0.1104 (120) S=0.230	0.1845 (120) S=0.044	0.0276 (120) S=0.765	0.1836 (120) S=0.045	0.1772 (120) S=0.053	0.0619 (120) S=0.502
MWT	-0.3249 (120) S=0.001	-0.0401 (120) S=0.664	-0.0481 (120) S=0.602	-0.2519 (120) S=0.006	-0.3120 (120) S=0.001	-0.1651 (120) S=0.072
	PEARSON CORRELATION			COEFFICIENTS		
	RHC1	RHC2	RHC3	RHC4	RHC5	RHC6
AGE	0.1658 (120) S=0.070	0.1000 (120) S=0.277	0.0725 (120) S=0.431	0.1824 (120) S=0.046	0.1780 (120) S=0.052	0.1299 (120) S=0.157
MWT	-0.2300 (120) S=0.011	-0.1380 (120) S=0.133	-0.0258 (120) S=0.780	-0.2322 (120) S=0.011	-0.2806 (120) S=0.002	-0.1827 (120) S=0.046
	PEARSON CORRELATION			COEFFICIENTS		
	RHD1	RHD2	RHD3	RHD4	RHD5	RHD6
AGE	0.0954 (120) S=0.290	0.2264 (120) S=0.013	0.2264 (120) S=0.013	0.2264 (120) S=0.013	0.2264 (120) S=0.013	0.2264 (120) S=0.013
MWT	-0.2117 (120) S=0.020	-0.2153 (120) S=0.018	-0.2153 (120) S=0.018	-0.2153 (120) S=0.018	-0.2153 (120) S=0.018	-0.2153 (120) S=0.018

Tabelle VI.5b. Korrelationen zwischen den Variablen Alter sowie Intelligenz und den Rohdaten der Testformen D, E und AW

	ROHD1		ROHD2		ROHD3		ROHD4		PEARSON		CORRELATION		COEFFICIENT	
AGE	0.1264 (120) \$0.189	0.0798 (120) \$0.386	0.0811 (120) \$0.379	0.1767 (120) \$0.053	0.2237 (120) \$0.014	0.1095 (120) \$0.234	0.2040 (120) \$0.025	0.3125 (120) \$0.001	0.1968 (120) \$0.031					
WMT	0.2589 (120) \$0.004	0.1364 (120) \$0.137	0.1014 (120) \$0.271	0.2061 (120) \$0.024	0.2942 (120) \$0.001	0.1901 (120) \$0.038	0.2901 (120) \$0.001	0.1570 (120) \$0.087	0.2120 (120) \$0.020					
8										PEARSON CORRELATION				
10	ROHE1		ROHE2		ROHE3		ROHE4		ROHE5		ROHE6		ROHE7	
AGE	0.3442 (120) \$0.001	0.1445 (120) \$0.115	0.0059 (120) \$0.1949	0.0749 (120) \$0.416	0.1532 (120) \$0.095	0.1547 (120) \$0.092	0.2159 (120) \$0.018	0.2414 (120) \$0.008	0.1411 (120) \$0.124					
WMT	0.2785 (120) \$0.002	0.0955 (120) \$0.299	0.0265 (120) \$0.1774	0.2636 (120) \$0.004	0.2211 (120) \$0.015	0.1529 (120) \$0.095	0.3107 (120) \$0.001	0.0779 (120) \$0.398	0.1578 (120) \$0.085					
18										PEARSON CORRELATION				
10	ROHAW1		ROHAW2		ROHAW3		ROHAW4		ROHAW5		ROHAW6		ROHAW7	
AGE	0.1607 (120) \$0.079	0.1097 (120) \$0.233	0.0926 (120) \$0.314	0.1726 (120) \$0.059	0.2638 (120) \$0.004	0.2659 (120) \$0.003	0.3125 (120) \$0.001	0.0052 (120) \$0.955	0.0817 (120) \$0.375					
WMT	0.1962 (120) \$0.032	0.1655 (120) \$0.071	0.0169 (120) \$0.855	0.3039 (120) \$0.001	0.3053 (120) \$0.001	0.1293 (120) \$0.180	0.2682 (120) \$0.003	0.1533 (120) \$0.095	0.1043 (120) \$0.257					
18										PEARSON CORRELATION				

4. Positionseffekte

Eine Untersuchung auf Positionseffekte setzt sich zum Ziel, herauszufinden, ob die Position, die eine jeweilige Parallelttestform (A-E) in der Rangfolge ihrer Darbietung einnimmt, Einfluß hat auf das betreffende Testergebnis.

In der nachfolgend dargestellten Varianzanalyse sind mit Hilfe des F-Wertes (F_{ratio}) und seiner zugehörigen Wahrscheinlichkeit (F_{prob}) die möglichen Positionseffekte untersucht worden.

Dabei ist:

$F \leq 0,01$ als Positionseffekt,

$F > 0,01$ als kein oder zufälliger Rangeffekt zu interpretieren.

Die Darstellung aller Werte von F_{ratio} und F_{prob} findet sich in Tabelle VI.6 auf der nächsten Seite.

Tabelle VI.6

SKT Form	SUBTEST								
	1	2	3	4	5	6	7	8	9
A	F ratio	0,643	3,977	0,208	1,958	1,741	1,765	3,421	4,976 0,899
	F prob	0,6332	0,0046	0,9335	0,1057	0,1458	0,1408	0,0111	0,0010 0,4671
B	F ratio	1,413	2,800	0,547	0,688	0,311	1,778	1,479	2,291 2,367
	F prob	0,2341	0,0292	0,7015	0,6018	0,8703	0,1381	0,2131	0,0638 0,0568
C	F ratio	0,572	0,379	0,118	1,091	0,918	0,929	1,345	0,409 0,301
	F prob	0,6837	0,8233	0,9758	0,3646	0,4563	0,4497	0,2576	0,8021 0,8766
D	F ratio	1,735	2,031	0,258	1,999	1,217	1,004	0,347	1,541 1,304
	F prob	0,1471	0,0946	0,9039	0,0993	0,3075	0,4085	0,8460	0,1951 0,2728
E	F ratio	0,405	2,287	0,755	1,403	1,641	1,109	1,952	4,622 0,288
	F prob	0,8046	0,0642	0,5569	0,2375	0,1754	0,3557	0,1066	0,0017 0,8854

Varianzanalyse zum Nachweis von Positionseffekten ($F \leq 0,01$) bei der Darbietung der einzelnen Testformen A bis E. Signifikante Werte finden sich nur bei Testform A, Subtest 2 und 8 sowie bei Testform E, Subtest 8.

VII Diskussion

Wie aus Tabelle VI.1 zu ersehen ist, liefert die Berechnung der Reliabilität unter der Annahme der strengen Parallelität für die Gruppe $n = 120$ (= die Gesamtzahl der getesteten Patienten aus drei Kliniken) ein insgesamt befriedigendes Resultat. Hervorzuheben sind die Reliabilitätskoeffizienten von Subtest 6 mit 0,820 und Subtest 7 mit 0,783, die als gut bezeichnet werden können. Negative Ausnahmen stellen die Subtests 2, 8 und 9 dar, jene items, die Gedächtnisfunktionen messen. Dieses Ergebnis scheint uns im wesentlichen darin begründet, daß jene drei Subtests für die betreffende Stichprobe von Patienten zu leicht waren, die Ergebnisse daher nur wenig streuten. Die Richtigkeit der Vermutung wird bestätigt, wenn man die Mittelwerte (M) und Standardabweichungen (SD) in Tabelle V.5 betrachtet. Die tatsächlich geringe Streuung der Ergebnisse in den Subtests 2, 8 und 9 bedeutet, daß alle Probanden eine ähnliche oder sogar fast gleiche Leistung erbracht haben, so daß dadurch keine systematische interindividuelle Varianz auftreten konnte. Dies wäre ein Fehler der Testkonstruktion, wenn das Verfahren für "Normale" oder Patienten ohne Funktionspsychose erstellt worden wäre. Im vorliegenden Falle aber dient es ja zur Erfassung von psychischen Leistungsstörungen bei reversiblen körperlich begründbaren Psychosen. An diesen Kranken wiederum vermag die Testreliabilität nicht überprüft zu werden, weil die Merkmale (= die Symptome) nicht konstant sind, in ihrer (quantitativen) Ausprägung also zu- oder abnehmen. Es genügt, wenn das Verfahren auf diese Weise zwischen Kranken (interindividuell) und zwischen dem Ausprägungsgrad der Krank-

heit (intraindividuell) zu differenzieren vermag. Rein testtheoretisch ist es allerdings tatsächlich so -dies gilt für den (psychometrischen) Normalbereich = für die Gesunden-, daß, wenn Tests offenkundig zu leicht oder zu schwer sind, die Probanden einer Stichprobe nicht mehr hinreichend unterschieden werden können. Die niedrige Reliabilität würde in diesem Falle weiter darauf hinweisen, daß nicht die "wahren Ergebnisse", die true-scores, sondern Störeinflüsse, die error-scores, die Variation bedingen.

Eine Bestätigung findet unsere Interpretation in dem Umstand, daß bei der Stichprobe $n = 76$, für die das Vorliegen eines Durchgangs-Syndroms ausgeschlossen wurde, die Reliabilitätskoeffizienten bei gleichbleibenden relativen Größenverhältnissen in der Regel noch niedriger liegen. Die niedrigsten Werte werden auch hier wieder bei den Gedächtnistests erzielt. Bei Subtest 2 tritt sogar ein negativer Wert auf. Der ist allerdings so gering, daß sich Interpretationsversuche (z.B. systematische Varianz gegen 0? Verhältnis der error-score- zur true-score-Varianz zugunsten der zuletzt genannten?) erübrigen.

Zur Deutung der niedrigen Reliabilitäten in den entspr. Subtests bieten sich noch die folgenden Überlegungen an: Die Auswahl der Patienten der Gruppe $n = 76$ ist auf Grund der Ergebnisse der Testform A vorgenommen worden. Die Wiederholungstestungen AW zeigen jedoch grössere Ergebnisfluktuationen, als anzunehmen war. Wenn jemand in der Testform A weniger als 4 Punkte erzielt hatte, so vermochte er anscheinend bei der Testwiederholung AW mehr als 4 Punkte zu erzielen. Er fiel somit aus dem vorgegebenen Bereich heraus. Eine modifizierte Normierung des Tests dergestalt, daß der Bereich, in dem kein Durchgangs-Syndrom angenommen wird, etwas größer gewählt ist, beispielsweise auf bis zu 6 Punkte,

würde die Reliabilitäten erhöhen. Die insgesamt deutlich geringere Reliabilität der Gruppe $n = 76$ gegenüber der Gruppe $n = 120$ läßt sich auch durch die normierungsbedingten höheren Ergebnisse vieler Probanden im Wiederholungstest erklären. Eine Betrachtung der Mittelwerte, der Varianzen und der Wertpunkte legt allerdings die Vermutung nahe, daß das weiter vorn angeführte Argument, die betr. Testaufgaben wären zu leicht, um zu differenzieren, eher als Erklärung in Frage kommt.

Die Reliabilitätsberechnungen unter der Annahme der nominellen Parallelität liefern zwei Koeffizienten, einen für Testform A und einen für die Testwiederholung AW. Da die Voraussetzungen der nominellen Parallelität leichter zu erfüllen sind als die der strengen Parallelität, wären dort insgesamt höhere Reliabilitätskoeffizienten zu erwarten. Wie ein Vergleich der Tabellen VI.1 und VI.2 zeigt, ist dies auch der Fall. Bei der Gruppe $n = 120$ heben sich ebenfalls die niedrigen Werte für die Gedächtnistests A_2 und A_8 heraus. Die Gründe dafür wurden schon besprochen. Im allgemeinen unterscheiden sich die beiden jeweiligen Reliabilitätskoeffizienten bis auf die Gedächtnisitems kaum.

Man kann also feststellen, daß jene Aufgaben, die nicht Gedächtnis messen, eine relativ hohe Reliabilität aufweisen. Dies gilt zumindest für die Gruppe $n = 120$, bei der sich die Reliabilitätskoeffizienten in einer Größenordnung von 0,522 bis 0,829 bewegen, während bei der Gruppe $n = 76$ die Werte doch bis 0,360 hinabreichen. Die Gedächtnisaufgaben sind anscheinend stärker anfällig gegen Störeinflüsse, wie man an den Reliabilitäten für Subtest A_2 und A_8 sowie AW_2 und AW_8 erkennt. Dies könnte dadurch erklärt werden, daß bei der ersten Darbietung des Tests die interindividuellen Unterschiede bei den Aufgaben voll zum tragen gekommen sind, bei der

zweiten Darbietung hingegen, wie man an der gesunkenen Reliabilität sieht, die Fehlerquote und auch die Streuung der Werte durch Trainings- und Erinnerungseinflüsse geringer waren. Bei den Gedächtnisaufgaben können also auch stärkere Wiederholungseinflüsse vorliegen. Die geringeren Reliabilitäten im Falle der kleineren Stichprobe wären dann durch die gleichen Argumente zu erklären.

Die Abschätzung der Reliabilitäten durch den Koeffizienten Alpha nach L.J. CRONBACH, der sich auf den Gesamttest bezieht, ergibt, wie Tabelle VI.3 zeigt, für die Gruppe $n = 120$ mit Werten um 0,8 zufriedenstellende Ergebnisse, während bei der Stichprobe $n = 76$ die Reliabilitäten wiederum niedriger liegen. Dies war nach den vorhergehenden Ergebnissen zu erwarten, da ja die einzelnen Subtests mit ihren Reliabilitäten ins Gesamtergebnis mit eingehen.

Unter bestimmten Bedingungen können auch die Korrelationen der entsprechenden Werte von Ersttest und Testwiederholung als Reliabilitäten interpretiert werden. Es zeigt sich ein ähnliches Ergebnis wie bei den Reliabilitätsberechnungen unter der Annahme der strengen und der nominellen Parallelität. Aus den in Tab. VI.4 dargestellten Werten ist zu entnehmen, daß hier die Gedächtnistests wieder etwas abgefallen sind. Die Reliabilitäten der Subtests 6 und 7 erreichen hingegen sehr hohe Werte. Die Korrelationen stützen also das oben erläuterte Ergebnis.

Wie aus Tabelle VI.5 und Abbildung VI.1a hervorgeht, machen sich Alterseinflüsse auf die Rohwerte bemerkbar. Sie sind jedoch nicht systematisch an einen Subtest oder an eine spezifische Klasse von Subtests gebunden. Es handelt sich am ehesten um Zufallseinflüsse. Die Absolutwerte der Korrelationskoeffizienten sind mit Größen um maximal 0,3 (also bei 9% gemeinsamer Varianz zwischen Alter und entsprechenden

Subtests) relativ gering. Der Einfluß des Alters auf das Ergebnis ist hier also deutlich nachweisbar, wenn auch gering.

Die Berechnungen der Korrelation zwischen der Variable "Intelligenz" und den Ergebnissen in den einzelnen Subtests haben folgendes ergeben: Intelligenz korreliert nicht mit den Resultaten in den Gedächtnisaufgaben (Subtest 2, 8 und 9), auch nicht mit denen im Zahlenlesen und Symbole zählen (Subtest 3 und 6). Diese Aufgaben scheinen also gegenüber den Einflüssen der Intelligenz relativ stabil zu sein. Die Intelligenz wirkt sich hingegen -Subtest 6 hier ausgenommen- auf die Geschwindigkeitsaufgaben aus, sowie auf das Gesamttest-Ergebnis. Dies mag daran liegen, daß für Probanden mit einem höheren Intelligenzniveau im Vergleich zu solchen mit niedrigerem Begabungsstand Testanweisungen und Testaufgaben allgemein adäquater sind.

Es wäre in diesem Zusammenhang interessant, den SKT als Intelligenztest -sinnvollerweise in unterdurchschnittlichen Bereichen- einzusetzen und die Ergebnisse daraufhin zu überprüfen.

Noch einmal: Intelligenzeinflüsse liegen beim SKT also fast durchgängig im Hinblick auf die Geschwindigkeitsaufgaben vor, hingegen nicht in Bezug auf die Gedächtnisaufgaben. Alterseinflüsse sind nicht eindeutig nachweisbar.

Bei der Frage, ob die Reihenfolge der Testformen bei der Abnahme der 5 Gesamttests eine Rolle spielt, zeigt sich (vgl. Tab. VI.6) ein interpretierbarer Positionseffekt nur in Testform A für die Subtests 2 und 8 (Gegenstände unmittelbar und mittelbar reproduzieren) sowie in Testform E wiederum für den Subtest 8. Da diese drei überzufälligen Zusammenhänge das Ergebnis von insgesamt einhundert entsprechenden Überprüfungen sind, dürfte es sich dabei um zufällige

Signifikanzen handeln, die etwa durch Versuchsleiterfehler oder durch eine ungünstige Testsituation zu erklären wären. Ein Positionseffekt, veranlaßt durch eine entsprechende Reihenfolge bei der Abnahme der einzelnen Testformen A bis E, läßt sich somit nicht nachweisen.

VIII Zusammenfassung

Die Überprüfung der Form A des Syndrom-Kurztests (SKT) auf seine Retestreliabilität an einer Untersuchungsstichprobe von 120 Patienten ergibt für die Annahme der strengen Parallelität gute Ergebnisse, für Subtest 6 z.B. einen Reliabilitätskoeffizienten von 0,820. Für die Annahme der nominellen Parallelität, die weniger anspruchsvolle Bedingungen fordert, liegen die entsprechenden Werte noch etwas höher, z.B. 0,827 für Subtest 6. Im Falle der Unterstichprobe $n = 76$ - sie besteht aus jenen Probanden, bei denen ein Durchgangssyndrom sicher ausgeschlossen werden konnte - sind die Ergebnisse etwas weniger befriedigend, was offensichtlich daran liegt, daß die Tests zu leicht sind, um unter Gesunden hinreichend zu differenzieren.

Für jene Subtests, die Gedächtnisfunktionen messen, ergeben sich erheblich niedrigere Reliabilitätskoeffizienten. Dies dürfte wiederum daran liegen, daß die Tests allgemein zu leicht sind und somit stärkere Wiederholungseinflüsse auftreten. Die Reliabilitätsbestimmung mit Hilfe der Korrelationen stützt das dargestellte Ergebnis. So beträgt der Reliabilitätskoeffizient für Subtest 6 wiederum 0,826. Die Abschätzung der Reliabilitäten nach L.J. CRONBACH für den Gesamttest, also über alle 9 Subtests hinweg, ergibt gute Resultate, so z.B. für die Gruppe $n = 120$ der Testform A einen Reliabilitätskoeffizienten von 0,840. Für die Teilstichprobe $n = 76$ fallen die Werte wieder etwas ab. Das Alter der Probanden übt nur einen geringen, nicht eindeutig nachweisbaren Einfluß auf die Meßwerte aus. Das Intelligenzniveau macht sich bei der Lösung der Geschwindigkeits-, nicht aber bei der Lö-

sung der Gedächtnisaufgaben bemerkbar. Ein Positionseffekt, gründend auf die entsprechende Reihenfolge in der Darbietung der Testformen A bis E, läßt sich nicht nachweisen.

IX Literaturverzeichnis

- 1 ARNOLD, K.: Beginn der Anwendung psychologischer Testmethoden in der Psychiatrie. In: Messung und Meßverfahren in der Psychopathologie. Reihe Psychopathometrie, Vless: Vaterstetten-München 1979
- 2 BÖCKER, F.: Eine Methode zur genaueren Erfassung von Bewußtseinstörungen und Durchgangs-Syndromen. Schweiz. Arch. Neurol. Psychiat. 88, 332-338 (1961)
- 3 BÖCKER, F., W. KINZEL: Durchführung und Auswertung des Syndromtests zur Bestimmung der Schwere von Funktionspsychosen. In: Das Ärztliche Gespräch 11, Tropon: Köln 1969
- 4 BONHOEFFER, K.: Die symptomatischen Psychosen im Gefolge von akuten Infektionen und inneren Erkrankungen. Deuticke: Leipzig-Wien 1910
- 5 BONHOEFFER, K.: Die exogenen Reaktionstypen. Arch. Psychiat. Nervenkr. 58, 58-70 (1917)
- 6 BRINKMANN, R., D.v. CRAMON, H. SCHULZ: The Munich Coma Scale (MCS). I. Neurol. Neurosurg. Psychiat. 39, 788-793 (1976)
- 7 DRUSCHKY, K.-F.: Orientierungsstörungen bei Funktionspsychosen. Diss., Erlangen 1971
- 8 DRUSCHKY, K.-F., W. KINZEL: Orientierungsstörungen bei Funktionspsychosen. Fortschr. Neurol. Psychiat. 43, 619-630 (1975)
- 9 ERZIGKEIT, H.: Schweregradmessungen bei Funktionspsychosen. 12. Fortbildungstagung für Krankenschwestern, -pfleger und Sozialarbeiter an den bayerischen Bezirks-Krankenhäusern und Universitäts-Nervenkliniken im Bezirks-Krankenhaus Mainkofen/Niederbayern am 10. und 17. Mai 1975

- 10 ERZIGKEIT, H.: Manual zum Syndrom-Kurztest Formen A - E. Vless: Vaterstetten-München 1977
- 11 ERZIGKEIT, H.: Das Durchgangs-Syndrom im Kurztest nach Erzigkeit. In: Ärztliche Praxis XXX Nr. 15, 429-430 (1978)
- 12 ERZIGKEIT, H.: Operationalisierungen zum Konzept "Durchgangs-Syndrom" - "Funktionspsychose". In: Psycho 4, 460-463 (1978)
- 13 ERZIGKEIT, H.: Syndrom-Kurztest. In: Messung und Meßverfahren in der Psychopathologie. Reihe Psychopathometrie Vless: Vaterstetten-München 1979
- 14 FISCHER, G.: Einführung in die Theorie psychologischer Tests. Huber: Bern 1974
- 15 FRIESSEM, D.H.: Kurzlehrbuch Nervenheilkundliches Stoffgebiet, Bd. 2: Psychiatrie. Freytag & Müller: München 1979
- 16 FUCHS, H.H.: Empirische Versuche mit verschiedenen Leistungsverfahren für Funktionspsychose-Testungen. Diss., Erlangen 1978
- 17 HEERKLOTZ, B.: Die Erforschung der Funktionspsychosen und die Entwicklung der Psychopathometrie. In: Messung und Meßverfahren in der Psychopathologie. Reihe Psychopathometrie. Vless: Vaterstetten-München 1979
- 18 HULL, C.H., N.H. NIE, J.G. JENKINS, K. STEINBRENNER, D.H. BENT: SPSS Second Edition. Mc Graw-Hill: New York 1975
- 19 HULL, C.H., N.H. NIE: SPSS Update. Mc Graw-Hill: New York 1979
- 20 KINZEL, W.: Die Trennung reversibler und irreversibler Syndrome körperlich begründbarer Psychosen. Phänomenologie, Analyse, experimentelle Ergebnisse und Bedeutung einer Syndromdifferenzierung. Nervenarzt 42, 585-590 (1971)
- 21 KINZEL, W.: Die Unterscheidung vorübergehender seelischer Störungen und bleibender seelischer Schäden. Methode zur Differenzierung körperlich begründbarer Psychosen. Münch. med. Wschr. 113, 117-123 (1971)

- 22 KINZEL, W.: Eine Verfahrensweise zur Trennung rückbildungs-
fähiger und nicht mehr rückbildungsfähiger seelischer Ver-
änderungen körperlich begründbarer Psychosen. Psychiat.
clin. 5, 367-379 (1972)
- 23 KINZEL, W.: Die Bedeutung der Psychopathometrie für die
moderne Psychiatrie. Eine erweiterte Begriffsbestimmung.
Psychopathometrie 1, 9-19, 1975
- 24 KINZEL, W.: Die Notwendigkeit der Unterscheidung von Funk-
tionspsychose und Defektsyndrom. In: Messung und Meßverfahren
in der Psychopathologie, hrsg. von ERZIGKEIT, H., S. LEHRL,
L. BLAHA, B. HEERKLOTZ. Vless: Vaterstetten-München 1978
- 25 KINZEL, W., J. KRAUS: Psychopathometrie bei endogenen
Psychosen. Abgrenzung der Schizophrenie und der Zykllothymie
vom organischen Defektsyndrom nach Hirnkontusion. Fortschr.
Neurol. Psychiat. 42, 113-156 (1974)
- 26 LEHRL, S.: Psychopathometrie in der Praxis. Z. Allg. ~
Med. 53, 1870-1880 (1977a)
- 27 LEHRL, S.: Manual zum MWT-B. Straube: Erlangen 1977
- 28 LEHRL, S.: Psychopathometrische Kurztests. Psycho 4,
282-291 (1978)
- 29 LEHRL, S.: Der MWT-B. In: Messung und Meßverfahren in der
Psychopathologie. Reihe Psychopathometrie. Vless: Vater-
stetten-München 1979
- 30 LEHRL, S., H. DAUN, R. SCHMIDT: Eine Abwandlung des HAWIE-
-Wortschatztests als Kurztest zur Messung der Intelligenz Er-
wachsener. Arch. Psychiat. Nervenkr. 214, 353-364 (1971)
- 31 LEHRL, S., J. MERZ, H. ERZIGKEIT, V. GALSTER: Der MWT-A, ein
wiederholbarer Intelligenz-Kurztest, der weitgehend unab-
hängig von seelisch-geistigen Störungen ist. Nervenarzt 45,
364-369 (1974)
- 32 LEHRL, S., J.V. GALSTER: Psychopathometrie und homogene
Syndromdynamik. Schweiz. Arch. Neurol. Psychiat. 118,
167-174 (1976)

- 33 LEHRL, S., H. ERZIGKEIT: Psychopathometrische Verfahren bei der Prüfung von Psychopharmaka. In: International Pharmacopsychiatry 12, 25-37 (1977)
- 34 LEHRL, S., H.H. FUCHS, J. LUGAUER, H. SCHUMACHER, G. NUSKO: Manual zur Funktionspsychose-Skala B. Vless: Vaterstetten-München 1977
- 35 LIENERT, G.A.: Testaufbau und Testanalyse. 2. Aufl. Beltz: Weinheim 1966
- 36 LUGAUER, J., S. LEHRL, H.H. FUCHS: Messung von Funktionspsychosen. Diagnostik 10, 241-245 (1977)
- 37 MAGNUSSON, D.: Testtheorie. F. Deuticke: Wien 1969
- 38 MERZ, J., S. LEHRL, V. GALSTER, H. ERZIGKEIT: MWT-B - ein Intelligenz-Kurztest. Psychiat. Neurol. med. Psychol. 27, 423-428 (1975)
- 39 MICHEL, L.: Grundlagen der Testdiagnostik: In: Handbuch der Psychologie, Bd. 6 (hrsg. von R. HEISS, K.J. GROFFMANN, L. MICHEL) Hogrefe: Göttingen 1964
- 40 RÖTH, F.G.: Experimentelle Untersuchungen über Intelligenzstörungen im Durchgangs-Syndrom. Arch. Psychiat. Nervenkr. 214, 127-130 (1971)
- 41 WEBER, E.: Grundriß der biologischen Statistik. 6. Aufl. Fischer: Stuttgart 1967
- 42 WIECK, H.H.: Zur Klinik der sogenannten symptomatischen Psychosen. Dtsch. med. Wschr. 81, 1345-1349 (1956)
- 43 WIECK, H.H.: Zum Begriff der Syndromgenese. Med. Welt 1961: 1183-1185
- 44 WIECK, H.H.: Grundzüge einer Syndromlehre der Funktionspsychosen. Verh. Dtsch. Ges. inn. Med. 75, 814-819 (1969)
- 45 WIECK, H.H.: Die Lehre von den Funktionspsychosen. Med. Welt 20 (N.F.): 936-941 (1969)
- 46 WIECK, H.H.: Zur Psychopathometrie. In: Angewandte Psychopathometrie, hrsg. von H.H. Wieck, Janssen-Symposien Bd. 8, Düsseldorf 1973

- 47 WIECK, H.H.: Neurologie und Psychiatrie in der Praxis.
 Schattauer: Stuttgart-New York 1974
- 48 WIECK, H.H.: Lehrbuch der Psychiatrie. Schattauer:
 Stuttgart-New York 1977
- 49 WIECK, H.H., F. BÖCKER: Besondere klinische Erscheinungs-
 weisen des Durchgangs-Syndroms. Med. Welt 1963: 1103-1108
- 50 WIECK, H.H., K. STÄCKER: Zur Dynamik des "amnestischen"
 Durchgangs-Syndroms. Arch. Psychiat. Nervenkr. 205,
 479-512 (1964)
- 51 WIECK, H.H., H. Erzigkeit: Zur Anwendung von psychopatho-
 metrischen Kurztests im klinischen Bereich. In: Der
 Krankenhausarzt 51, 31-36 (1978)

Hiermit möchte ich mich bei Herrn Professor Dr. H.H. WIECK, der die Anregung zu dieser Arbeit gab, und bei Herrn Professor Dr. W. KINZEL, der mir das Thema überließ, für die Ermöglichung dieser Arbeit bedanken.

Weiterhin möchte ich den Herrn Professoren Dr. F.P. GALL und Dr. W. SCHIEFER Dank sagen. Neben Herrn Professor Dr. H.H. WIECK gestatteten auch sie mir, ihre Patienten zu untersuchen.

Mein besonderer Dank gilt Herrn Dipl. Psych. Dr. H. ERZIGKEIT für die vielen wertvollen Anregungen und Herrn Dipl. Psych. K. ARNOLD für die tatkräftige Unterstützung bei der Auswertung der Untersuchungen.

Lebenslauf

Geboren am 23. September 1950 in Nürnberg als Sohn des Justizbeamten Armin Salomo und seiner Ehefrau Gertrud, geborene Oßmann. Eintritt in die Volksschule im Herbst 1956, Übertritt an das Hans-Sachs-Gymnasium in Nürnberg im Herbst 1960, Ablegung des Abiturs am 20. Juni 1970. Beginn des Studiums der Werkstoffwissenschaften zum Wintersemester 1970/71 an der Friedrich-Alexander Universität Erlangen. Immatrikulation im Fachbereich Medizin zum Wintersemester 1975/76. Ablegung der Vorprüfung 1977, Beendigung des Studiums mit dem dritten Teil der ärztlichen Prüfung am 5. Mai 1982. Erteilung der Approbation als Arzt am 14. Mai 1982. Am 16. Oktober 1981 Eheschließung mit meiner Ehefrau Shatha, geborene Yassin. Aufnahme der Tätigkeit als Assistenzarzt an der 4. Medizinischen Klinik des Nürnberger Krankenhauses zum 1. Juni 1982. Einberufung zur Bundeswehr als grundwehrdienstleistender Arzt am 6. September 1982.

